



Risques de taux et de longévité: Modélisation dynamique et Applications aux produits dérivés et à l'assurance-vie

Harry Bensusan

► To cite this version:

Harry Bensusan. Risques de taux et de longévité: Modélisation dynamique et Applications aux produits dérivés et à l'assurance-vie. Probabilités [math.PR]. Ecole Polytechnique X, 2010. Français. NNT: . pastel-00563792

HAL Id: pastel-00563792

<https://pastel.archives-ouvertes.fr/pastel-00563792>

Submitted on 7 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT DE L'ÉCOLE POLYTECHNIQUE

Spécialité :
MATHÉMATIQUES APPLIQUÉES

Présentée par

Harry BENSUSAN

pour obtenir le grade de DOCTEUR DE L'ÉCOLE POLYTECHNIQUE

Sujet :

**Risques de taux et de longévité : Modélisation dynamique et applications
aux produits dérivés et à l'assurance vie**

**Interest rate and longevity risks : dynamic modelling and applications to
derivative products and life insurance**

Soutenue le 22 Décembre 2010 devant le jury composé de :

Nicole EL KAROUI	<i>Directrice de thèse</i>
Jean Pierre FOUQUE	<i>Rapporteur</i>
Christian GOURIEROUX	<i>Rapporteur</i>
Sylvie MÉLÉARD	<i>Examineur</i>
Christophe MICHEL	<i>Examineur</i>
Nizar TOUZI	<i>Examineur</i>
Alain TROGNON	<i>Président</i>

Remerciements

En premier lieu, je tiens à remercier de tout cœur ma directrice de thèse, Nicole El Karoui, pour la qualité exceptionnelle de son encadrement. Son savoir, son dynamisme et sa rigueur ont fortement contribué à la rédaction de cette thèse. C'était un grand honneur d'avoir découvert le monde de la recherche à ses côtés : elle m'a donné le goût de la recherche en me transmettant sa passion. J'admire son savoir immense mais je pense que j'admire encore plus ses qualités humaines. Nicole a toujours été disponible pour moi et a répondu à toutes mes questions avec gentillesse et enthousiasme. Je lui suis vraiment reconnaissant pour le soutien et la confiance qu'elle m'a accordé.

Mes remerciements respectueux à Jean Pierre Fouque et Christian Gourieroux de m'avoir fait l'honneur de rapporter ma thèse. Ils m'ont envoyé de nombreuses corrections et un ensemble de conseils précieux qui m'ont permis d'améliorer la rédaction.

Je présente mes sincères remerciements à Sylvie Méléard, Christophe Michel, Nizar Touzi et Alain Trognon de m'avoir fait l'honneur d'accepter de participer à mon jury de thèse.

Je tiens à remercier particulièrement Christophe Michel qui m'a accueilli pendant trois années dans son équipe de recherche du Crédit Agricole et qui m'a suivi, aidé et conseillé tout au long de cette thèse. Je remercie aussi Anas, membre de l'équipe, d'avoir partagé de nombreuses connaissances et d'avoir accepté de rédiger des articles ensemble. Je remercie Florence pour l'aide à la relecture de la thèse et je remercie vivement Vincent pour ses nombreux conseils avisés et son aide précieuse. Enfin, je remercie toute l'équipe de recherche de CA-CIB pour son accueil chaleureux, et plus particulièrement Vincent, Sophie, Christophe M. et Christophe Z., Kenza, Florence, Thomas, Eric avec qui j'ai pu tisser des liens d'amitié.

Mes remerciements d'adressent également aux membres du groupe de longévité et plus particulièrement Caroline Hillairet, Stéphane Loisel, Yahia Salhi, Pauline Barrieu, Isabelle Camillier et Trung Lap NGuyen. Ces réunions ont été très enrichissantes et nous avons pu travailler ensemble et partager des discussions sur de nombreux sujets passionnants.

Je remercie vivement Chi Tran Viet et Sylvie Méléard pour m'avoir aidé à me plonger dans le domaine de la dynamique de population. Chi m'a aidé à de nombreuses reprises en répondant à mes questions et en me fournissant des pistes de réflexion très intéressantes. Je remercie aussi les membres du groupe de travail de biologie de l'école polytechnique.

Un grand merci à Martino Grasseli et José Da Fonseca pour les nombreux conseils qu'ils m'ont donné sur les modèles de Wishart et pour l'intérêt qu'ils ont montré sur mes travaux de recherche.

Je pense aussi aux doctorants du CMAP, et en particulier à Gilles-Edouard, Emilie, Laurent, Isabelle, Mohammed et Trunglap ainsi qu'à l'ensemble des chercheurs du CMAP avec lesquels j'ai pu partager des moments privilégiés. Je pense en particulier à Peter, Emmanuel et Nizar qui ont été disponible et m'ont donné des conseils précieux. Merci au personnel administratif du CMAP, en particulier à Nassera, Anna, Sandra, Nathalie et Alexandra, pour leur gentillesse et leur aide dans leur bonne humeur. Je remercie aussi l'administrateur informatique Sylvain de son assistance ainsi qu'à Aldjia pour m'avoir aidé dans de nombreuses démarches. Enfin, je pense à l'ensemble des membres de l'école doctorale et en particulier Audrey et Fabrice qui m'ont aidé dans toutes les démarches.

Je voudrais remercier tous mes amis proches qui ont contribué à mon épanouissement. Merci à Dany, mon "colloc" adoré avec qui j'ai passé une année géniale, mes "petites sœurs" Cynthia et Karen et l'ensemble de mes amis proches Bastien, Matthieu, Jérôme, Ruth, Samuel, Yaniv, Muriel, Sara, Audrey, Claire, Jeremy, Elodie, Georges et tous les autres.

Une pensée particulière à mon ancien directeur d'école, David Elbaz, pour qui j'ai beaucoup d'affection et qui a participé à mon épanouissement de l'école primaire jusqu'au collège. Je pense aussi à l'ensemble des professeurs ainsi qu'à mes anciens camarades d'école.

Je tiens aussi à remercier mes cousins et cousines avec lesquels j'ai partagé des moments et des vacances inoubliables et en particuliers Sarah, Avner, Carole, Abraham et Béatrice. Un grand merci aux très nombreuses personnes de ma famille que je n'ai pas cité mais à qui je pense très fort.

Un affectueux remerciement à ma Eva pour sa patience, sa gentillesse et son soutien. Je suis ravi de l'avoir rencontré et j'espère partager encore de nombreux moments de bonheur à ses côtés. Petit coucou à ses adorables parents, à Jade sa petite sœur ainsi qu'aux nouveaux mariés Anaïs et Romain à qui je souhaite plein de bonheur.

Un remerciement particulier à mes grands frères, Franck, Dan, Teddy et Steeven ainsi qu'à mes belles sœurs Agathe et Nathalie, que j'aime profondément. Mes frères m'ont montré la voie à suivre, m'ont soutenu et m'ont réconforté tout au long de mes études. J'ai de la chance d'avoir des frères aussi attentionnés et je leur adresse ma profonde affection. J'en profite pour faire un coucou à mes adorables neveux et nièce, Alan et Eden, en espérant qu'ils gardent la même joie de vivre tout au long de leur vie.

Le mot de la fin est adressé à mes chers parents à qui je dédie cette thèse. Les mots que je pourrais écrire ne sont pas à la hauteur de ce que je peux ressentir à leur égard. Je les aime du fond du cœur et les remercie pour l'amour et le soutien qu'ils m'ont donné tout au long

de ma vie. Je les remercie surtout de m'avoir inculqué les valeurs morales qui m'ont permis de devenir la personne que je suis. Depuis mon enfance, je n'ai vécu que des moments de joie à leur côté et j'espère qu'on partagera tous ensemble de nombreuses années de bonheur.

Résumé

Cette thèse se divise en trois parties.

La première partie est constituée des chapitres 2 et 3 dans laquelle nous considérons des modèles qui décrivent l'évolution d'un sous-jacent dans le monde des actions ainsi que l'évolution des taux d'intérêt. Ces modèles, qui utilisent les processus de Wishart, appartiennent à la classe affine et généralisent les modèles de Heston multi-dimensionnels. Nous étudions les propriétés intrinsèques de ces modèles et nous nous intéressons à l'évaluation des options vanilles. Après avoir rappelé certaines méthodes d'évaluation, nous introduisons des méthodes d'approximation fournissant des formules fermées du smile asymptotique. Ces méthodes facilitent la procédure de calibration et permettent une analyse intéressante des paramètres.

La deuxième partie, du chapitre 4 au chapitre 6, étudie les risques de mortalité et de longévité. Nous rappelons tout d'abord les concepts généraux du risque de longévité et un ensemble de problématiques sous-jacentes à ce risque. Nous présentons ensuite un modèle de mortalité individuelle qui tient compte de l'âge et d'autres caractéristiques de l'individu qui sont explicatives de mortalité. Nous calibrons le modèle de mortalité et nous analysons l'influence des certaines caractéristiques individuelles. Enfin, nous introduisons un modèle microscopique de dynamique de population qui permet de modéliser l'évolution dans le temps d'une population structurée par âge et par traits. Chaque individu évolue dans le temps et est susceptible de donner naissance à un enfant, de changer de caractéristiques et de décéder. Ce modèle tient compte de l'évolution, éventuellement stochastique, des taux démographiques individuels dans le temps. Nous décrivons aussi un lien micro/macro qui fournit à ce modèle microscopique de bonnes propriétés macroscopiques.

La troisième partie, concernant les chapitres 7 et 8, s'intéresse aux applications des modélisations précédentes. La première application est une application démographique puisque le modèle microscopique de dynamique de population permet d'effectuer des projections démographiques de la population française. Nous mettons aussi en place une étude démographique du problème des retraites en analysant les solutions d'une politique d'immigration et d'une réforme sur l'âge de départ à la retraite. La deuxième application concerne l'étude des produits d'assurance-vie associant les risques de longévité et de taux d'intérêt qui ont été étudiés en détails dans les deux premières parties de la thèse. Nous nous intéressons tout d'abord à l'étude du risque de base qui est généré par l'hétérogénéité des portefeuilles de rentes. De plus, nous introduisons la *Life Nominal Chooser Swaption* (LNCS) qui est un produit de transfert de risque des produits d'assurance-vie : ce produit a une structure très intéressante et permet à une assurance détenant un portefeuille de rente de transférer intégralement son risque de taux d'intérêt à une banque.

Abstract

This thesis is divided into three parts :

The first part is constituted by chapter 2 and chapter 3 in which we consider models that describe the evolution of a stock and the evolution of interest rates. Using Wishart processes, these models belong to affine class and extend multidimensional Heston model. We study the intrinsic properties and we consider the evaluation of vanilla options. After having recalled evaluation methods, we introduce approximation methods that provide closed formula for the asymptotic smile. These methods make easier the calibration procedure and allow an analysis of the parameters.

The second part, from chapter 4 to chapter 6, studies mortality and longevity risks. First, we recall general concepts of longevity risk and the main challenges we have to tackle. Then, we study a model of mortality that takes into account age and other individual features that could explain mortality. This mortality model is calibrated on specific data what allows a study of the impact on mortality of different characteristics. Finally, we introduce a microscopic model for population dynamic in order to study the evolution in the future of an age-and-trait-structured population. Each person can be affected by different events : birth, death and evolution of his features. Moreover, this model takes into account of the stochastic evolution of the demographic rates in the time. We study a micro/macro link providing good macroscopic properties for this microscopic model.

The last part, constituted by chapter 7 and chapter 8, deals with applications of the previous models. The first application concerns demography : the association of the microscopic model for population dynamic and the mortality model provide demographic projections of the French population. We consider the retirement issue by analyzing the solutions of an immigration policy and a reform about the start age of retirement. The second application refers to insurance life products associating longevity and interest rate risks that have been detailed in the previous parts of the thesis. We study the basis risk generated by the heterogeneity of portfolio. Moreover, we introduce the *Life Nominal Chooser Swaption* (LNCS), a derivative product in insurance life for the transfer of financial risks. The design of the Life Nominal Chooser Swaption is very interesting allowing an assurance owning a portfolio to transfer his interest rate risk to a bank.

Avant Propos

Cette thèse est divisée en trois parties : la première partie, du chapitre 2 à 3, concerne l'étude de modèles affines à volatilité stochastique utilisant principalement le processus de Wishart. La deuxième partie, du chapitre 4 à 6, aborde une modélisation mathématique de la mortalité et de la dynamique de la population. La dernière partie (chapitre 7 à 9) concerne les applications en démographie et en finance de la deuxième partie en faisant référence aux travaux de la première partie.

L'organisation est faite comme suit :

- Le chapitre 1 est une présentation des différents problèmes abordés dans cette thèse ainsi que des différents résultats obtenus.
- Le chapitre 2 est une étude du "Wishart Volatility Model" pour la modélisation d'un sous-jacent dans le monde des actions. Le modèle est présenté avec ses propriétés intrinsèques (extension du Heston multidimensionnel, corrélation stochastique) et nous proposons une approximation du smile utilisant les méthodes de perturbation ainsi qu'une étude de sa robustesse. La méthode de perturbations régulières permet de trouver une approximation pour une volatilité dont l'échelle de temps est lente, et la méthode de perturbations singulières s'applique au cas d'une volatilité multi-échelle. Enfin, nous étudions l'influence des différents paramètres du modèle précédent.
- Le chapitre 3 est une adaptation du modèle précédent pour la modélisation des taux d'intérêt. Nous développons un modèle affine qui permet de prendre en compte de nombreux risques comme le risque de taux, le risque de volatilité et le risque de corrélation. De plus, nous exhibons un modèle dont le skew des options sur taux court est capté par un facteur du modèle, ce qui rend le risque de skew potentiellement couvrable. Enfin, nous étudions l'impact des paramètres du modèle sur la courbe de taux, les smiles de caplets et les surfaces de volatilité de swaption.
- Le chapitre 4 est une introduction aux problématiques liées au risque de longévité. Nous abordons les différents enjeux, les problèmes concernant la modélisation, la gestion de ce risque ainsi que le transfert de ce risque.
- Le chapitre 5 concerne la modélisation stochastique de la mortalité par âge et par traits. Une étude socioéconomique est présentée permettant de connaître les facteurs explicatifs de mortalité. En utilisant la base de données de l'Échantillon Démographique Permanent (EDP) de l'INSEE, nous verrons que nous pouvons créer de nouvelles tables de mortalité prospectives en fonction de l'âge mais aussi en fonction de différentes caractéristiques.
- Le chapitre 6 est une introduction à un modèle microscopique de dynamique de population pour une population structurée par âge et par traits. Cette modélisation probabiliste reprend des concepts utilisés en biologie que nous adaptons au cadre de

l'évolution d'une population d'êtres humains.

- Le chapitre 7 concerne les applications du modèle de mortalité en démographie, avec une analyse des scénarios d'évolution possibles dans le futur. Nous proposons une modélisation du problème des retraites d'un point de vue démographique. Nous analysons l'enjeu de l'immigration et de l'âge de la retraite au travers de notre modélisation.
- Le chapitre 8 aborde les applications financières du modèle de mortalité et de dynamique de populations. L'association de ces modèles fournit une modélisation utile pour les produits d'assurance vie avec une étude quantitative du risque de base. De plus, nous nous intéressons aux méthodes de transfert de risque de taux d'intérêt des produits d'assurance vie. Nous introduirons la problématique du transfert de risque ainsi que les difficultés majeures sous-jacentes. Enfin, nous mettons en place en détails un produit final que nous appelons la Life Nominal Chooser Swaption (LNCS) et nous l'étudions en détails. Ce produit permet à une assurance de transférer "au mieux" le risque de taux d'intérêt inhérent à ses portefeuilles de rentes.

Table des matières

1	Introduction	1
1.1	Étude du modèle de Wishart dans le monde des actions et dans le monde des taux d'intérêt	3
1.1.1	Étude du modèle de Wishart dans le monde des actions	4
1.1.2	Étude du modèle de Wishart dans le monde des taux d'intérêt	5
1.2	Risque de longévité et dynamique de population	7
1.2.1	Introduction à la problématique sur la longévité (Survey)	8
1.2.2	Modèle de mortalité par âge et par traits	9
1.2.3	Modèle microscopique de dynamique de population	10
1.3	Applications à la démographie et aux produits d'assurance vie	11
1.3.1	Prévisions démographiques et études du problème des retraites	12
1.3.2	Applications à la finance : étude du risque de base d'un portefeuille et structuration d'un produit	13
2	Wishart Stochastic Volatility	17
2.1	The Wishart Process	20
2.1.1	Definition of the Wishart processes	20
2.1.2	Examples	21
2.1.3	Determinant dynamics and change of probability measure	23
2.2	Wishart stochastic volatility model in the stock derivatives market	25
2.2.1	Presentation of the Wishart volatility model	25
2.2.2	Evaluation of options by Monte Carlo methods	31
2.2.3	Evaluation of options by FFT	33
2.3	Smile dynamics in the Wishart volatility model	35
2.3.1	Specification of the Wishart volatility model	36
2.3.2	Condition for the perturbation method	36
2.3.3	Perturbation Method of the Riccati equations	38
2.3.4	Smile dynamics for a slow volatility time scale	39
2.3.5	Smile dynamics for a two-scale volatility	43
2.3.6	Numerical Applications	48

TABLE DES MATIÈRES

2.4	Study of the influence of the parameters	57
2.4.1	Influence of the parameter β	58
2.4.2	Influence of the correlation R	58
2.4.3	Influence of the volatility of volatility Q	59
2.4.4	Influence of the initial volatility V_0	64
2.5	Conclusion	73
2.6	Appendix	74
3	Modèle de Wishart pour les taux d'intérêt	79
3.1	Présentation du modèle de Wishart sur les taux	82
3.1.1	Description du modèle	82
3.1.2	Évaluation du prix des zéro-coupon	84
3.1.3	Description de la probabilité forward neutre	88
3.1.4	Transformation de Laplace des facteurs	88
3.2	Méthodes d'évaluation des options vanilles	92
3.2.1	Évaluation des swaptions	92
3.2.2	Évaluation des caplets	96
3.3	Étude d'un modèle de Wishart équivalent à un modèle de Heston	97
3.3.1	Présentation du modèle	97
3.3.2	Smile de caplets et surface de volatilité à la monnaie (ATM)	99
3.3.3	Influence des paramètres du modèle	103
3.3.4	Étude d'un modèle Wishart en dimension 2	108
3.4	Propriétés intrinsèques du modèle de Wishart	109
3.4.1	Comparaison modèle Wishart/ modèle affine (Daï-Singleton)	109
3.4.2	Spécification d'un modèle qui contrôle la dynamique du skew	111
3.5	Étude du smile asymptotique	116
3.5.1	Méthode d'approximation de smile en temps petit	116
3.5.2	Méthode de perturbations régulières	117
3.6	Conclusion	120
3.7	Annexes	121
4	Understanding, Modelling and Managing Longevity Risk	131
4.1	Main characteristics of longevity risk	133
4.1.1	Age, period and cohort life tables and available mortality data	133
4.1.2	Observed heterogeneity, inter-age dependence and basis risk	135
4.2	Modelling longevity risk	136
4.2.1	Some standard models	136
4.2.2	Observed heterogeneity	139
4.2.3	Modelling inter-age dependence and inter-temporal correlations	142
4.3	New regulations, economic and sociologic impacts of longevity risk	143

TABLE DES MATIÈRES

4.3.1	Technical provisions	143
4.3.2	Solvency Capital Requirements	144
4.3.3	Risk margin	147
4.3.4	Impact of longevity risk on the economy	147
4.4	Transferring longevity risk	148
4.4.1	Convergence between insurance and capital markets	148
4.4.2	Recent developments in the transfer of longevity risk	149
4.4.3	Various longevity indices	150
4.4.4	q-forwards	151
4.4.5	Longevity swap transactions and basis risk	152
4.5	Modelling issues for pricing	153
4.5.1	Pricing methodologies	153
4.5.2	Dynamic modelling of the underlying risk	157
4.6	Conclusion	159
5	Modélisation de la mortalité individuelle	165
5.1	Étude du modèle de Cairns décrivant la mortalité par âge	168
5.1.1	Description et justification du modèle de mortalité	168
5.1.2	Calibration du modèle de mortalité et scénarios de projection	176
5.1.3	Backtesting du modèle de Cairns	179
5.1.4	Extension du modèle de Cairns à deux facteurs	183
5.1.5	Extension à un modèle prenant en compte les catastrophes	184
5.2	Étude des facteurs explicatifs de la mortalité	186
5.2.1	Sexe de l'individu	187
5.2.2	Statut matrimonial et vie familiale	188
5.2.3	Catégorie socioprofessionnelle	189
5.2.4	Niveau d'éducation	189
5.2.5	Niveau de revenu	190
5.2.6	Étude du déterminisme social et génétique	194
5.2.7	Dépendance des individus et prévention	195
5.2.8	Bilan sur les facteurs explicatifs de mortalité	196
5.3	Modélisation de la mortalité des individus par âge et par traits	197
5.3.1	Description du modèle	197
5.3.2	Méthode de calibration	199
5.4	Étude quantitative de l'influence des caractéristiques sur la mortalité	202
5.4.1	Analyse de l'Échantillon Démographique Permanent	202
5.4.2	Étude de l'influence de la catégorie socioprofessionnelle	208
5.4.3	Étude de l'influence du niveau de diplôme	214
5.4.4	Étude de l'influence du statut matrimonial	217
5.5	Conclusion	221

TABLE DES MATIÈRES

6	Modélisation microscopique de la dynamique de population	227
6.1	Introduction aux processus d'évolution de population	230
6.1.1	Processus ponctuels et processus de Poisson	231
6.1.2	Processus de naissance et mort	232
6.1.3	Processus ponctuel discret d'évolution	236
6.2	Description des paramètres d'évolution de la population	238
6.2.1	Intensités individuelles d'évolution	239
6.2.2	Évolution des intensités individuelles dans le temps	240
6.2.3	Hypothèses sur les intensités d'évolution	241
6.2.4	Probabilité de survie	242
6.2.5	Lien entre les notions de taux et d'intensité	243
6.3	Processus microscopique	244
6.3.1	L'espace des mesures ponctuelles	244
6.3.2	Définition du processus d'évolution	245
6.3.3	Simulation du processus d'évolution	248
6.3.4	Migration et immigration	250
6.3.5	Extensions possibles	252
6.4	Propriétés mathématiques du processus	253
6.4.1	Existence et unicité forte du processus	253
6.4.2	Propriétés de moments du processus	255
6.4.3	Études du générateur infinitésimal du processus Z_t	257
6.4.4	Dynamique du processus d'évolution	259
6.4.5	Distribution moyenne de la population	260
6.5	Renormalisation du processus microscopique	262
6.5.1	Définitions et propriétés mathématiques	263
6.5.2	Étude du processus limite pour les grandes populations	264
6.5.3	Approximation par un modèle macroscopique aléatoire	270
6.5.4	Interprétation de l'équation aux dérivées partielles stochastique . . .	272
6.5.5	Taux de reproduction net	273
6.6	Conclusion	274
6.7	Annexes	275
7	Projections démographiques et étude du problème des retraites	285
7.1	Cadre d'évolution d'une population française	287
7.1.1	Modèle de dynamique de population	287
7.1.2	Taux de mortalité	290
7.1.3	Taux de Natalité	293
7.1.4	Taux d'évolution de traits	298
7.1.5	Migration et immigration	300
7.2	Projections démographiques de la population française	304

TABLE DES MATIÈRES

7.2.1	Bilan démographique de la population française en 2007	304
7.2.2	Étude qualitative de l'évolution de la population française	306
7.2.3	Projections démographiques sans immigration	309
7.2.4	Projections démographiques avec prise en compte de l'immigration .	319
7.3	Problème des retraites et stratégies possibles	322
7.3.1	Introduction au problème des retraites	322
7.3.2	Impact des flux migratoires	323
7.3.3	Impact de l'âge de départ à la retraite	327
7.3.4	Analyse des solutions proposées	329
7.3.5	Formulation mathématique du problème des retraites	329
7.4	Conclusion	330
8	Risques financiers des produits d'assurance vie	337
8.1	Description de la modélisation du risque de longévité	340
8.1.1	Processus d'évolution	340
8.1.2	Lien Micro/Macro	341
8.1.3	Modélisation d'un taux de mortalité individuel	342
8.2	Étude du risque de longévité d'un portefeuille de rente	344
8.2.1	Évaluation du risque systémique de longévité	344
8.2.2	Étude de l'hétérogénéité du portefeuille	346
8.3	Introduction à la couverture de produits d'assurance vie	351
8.3.1	Compréhension et gestion des risques en assurance vie	351
8.3.2	Idées de produits de transfert du risque de taux	353
8.4	Longevity Nominal Chooser Swaption (LNCS)	359
8.4.1	Introduction au produit	359
8.4.2	Rappels de certains concepts financiers	364
8.4.3	Étude de la Longevity Nominal Chooser Swaption	368
8.4.4	Étude qualitative	372
8.4.5	Étude quantitative	374
8.4.6	Piste d'amélioration de la modélisation	382
8.5	Conclusion	384
8.6	Annexes	385

TABLE DES MATIÈRES

Chapitre 1

Introduction

Introduction

Cette thèse est constituée de trois parties. Dans la première partie, nous considérons la modélisation mathématique de certains risques financiers. Ces dernières années, la modélisation des marchés financiers a été un domaine de recherche effervescent mais le marché se complexifiant de plus en plus, il faut trouver de nouveaux modèles afin de décrire les phénomènes de marché. La modélisation que nous étudions s'appuie principalement sur les processus de Wishart et permet d'obtenir des modèles à volatilité stochastique qui ont des propriétés intéressantes. Tout d'abord, nous étudions en détails les propriétés du "Wishart Volatility Model" qui est une extension multidimensionnelle du modèle de Heston dans le domaine des actions. En utilisant la méthode des perturbations régulières et singulières, nous établissons des formules d'approximation de smiles facilitant la procédure de calibration du smile et permettant de comprendre l'influence des différents paramètres du modèle.

Nous proposons une nouvelle modélisation des taux d'intérêt avec un modèle affine décrivant le taux court comme combinaison linéaire de facteurs affines qui dépendent d'un processus de Wishart. Le caractère affine du modèle donne la possibilité de calculer efficacement les prix d'options vanilles constituées des caplets et des swaptions. Nous analysons l'impact des paramètres du modèle sur les différentes données de marché et nous décrivons un modèle particulier pour lequel le skew (défini dans la suite) des options sur taux court est capté par un facteur du modèle. Le risque lié à la dynamique du skew, qui est considéré comme exotique, pourrait éventuellement devenir couvrable en utilisant ce modèle. Ainsi, cette modélisation des taux d'intérêt se révèle assez efficace en tenant compte de certains phénomènes de marché.

Dans la deuxième partie, nous abordons la définition, la compréhension et la gestion des risques de mortalité et de longévité. L'évolution de la longévité des individus est considérée

comme stochastique ce qui constitue un risque important pour de nombreuses institutions que ce soit le gouvernement (avec des enjeux économiques comme le système des retraites) ou les assurances, réassurances et fonds de pension qui doivent gérer le risque de leur produits d'assurance vie. Ce sont les motivations qui nous ont poussées à chercher un modèle efficace décrivant la longévité des individus.

Inspirés par des travaux en mathématiques appliquées à la biologie et à l'écologie, nous introduisons un modèle microscopique de dynamique de population afin de décrire l'évolution d'une population d'êtres humains. La modélisation se fait au niveau individuel et chaque individu est caractérisé par son âge et par certaines variables appelées *traits*. Nous décrivons un modèle permettant de tenir compte de taux démographiques stochastiques. En effet, la mortalité des individus d'une population évolue de manière aléatoire il n'est pas raisonnable de considérer une évolution déterministe. Nous étudions la simulation de ce modèle ainsi que ses nombreuses propriétés mathématiques. Enfin, nous analysons le cas des grandes populations pour lequel nous pouvons déduire des approximations macroscopiques. Ces approximations vont généraliser certaines équations classiques en démographie, ce qui assure une cohérence du modèle dans le cas de grandes populations.

Pour une utilisation efficace de ce modèle, nous avons aussi modélisé la dynamique du taux de mortalité des individus de la population. En s'inspirant du modèle de Cairns qui est un modèle standard de mortalité stochastique par âge, nous avons proposé un modèle de mortalité par âge et par traits. Nous avons analysé quelles étaient les caractéristiques explicatives de mortalité et nous avons calibré le modèle sur les données de l'Échantillon Démographiques Permanent (EDP) de l'INSEE. Cela fournit une modélisation de la mortalité individuelle qui tient compte d'un ensemble d'informations sur chaque individu.

Dans la dernière partie, nous décrivons les applications des modélisations précédentes. La première application concerne une étude démographique ainsi qu'une introduction au problème des retraites. La conjoncture démographique actuelle suscite de nombreux débats concernant la politique à adopter dans les années à venir, en particulier pour le paiement des retraites. Nous verrons qu'il est possible de tenir compte de nombreux concepts associés à des stratégies politiques dans le modèle d'évolution comme le taux d'immigration et l'âge de départ à la retraite. En utilisant cette modélisation nous avons étudié les stratégies politiques concernant l'immigration et l'âge de départ à la retraite qui permettraient d'avoir une situation démographique "acceptable".

Bien que les domaines de la finance, de la biologie et de la démographie semblent assez éloignés les uns des autres, nous avons essayé de décrire une application connexe des parties précédentes dans le domaine de l'assurance vie. Les assurances possèdent des portefeuilles de rentes qui rendent compte de nombreux risques. Tout d'abord, nous définissons le risque

de base qui s'explique par la différence de mortalité qu'il existe entre la mortalité d'une population d'assurés et la mortalité d'une population nationale. Ensuite, nous utilisons le modèle de dynamique de population ainsi que le modèle de mortalité pour estimer le risque de base d'un portefeuille de rentes en tenant compte de l'ensemble de l'information disponible sur les individus.

De plus, les produits d'assurance vie comportent non seulement un risque lié à la longévité mais aussi certains risques financiers comme le risque de taux d'intérêt. Par exemple, l'évaluation des portefeuilles de rentes dépend d'une part de l'évolution de la longévité des individus et d'autre part de l'évolution des taux d'intérêt. Nous introduisons un nouveau produit Longevity Nominal Chooser Swaption (LNCS) permettant à l'assurance de transférer le risque financier inhérent aux produits d'assurance vie. En effet, l'assurance est spécialisée dans la gestion du risque de longévité mais éprouve des difficultés à gérer ses risques financiers tels que le risque d'évolution des taux d'intérêt. L'idée est de créer un produit qui permet à l'assurance de transférer ce risque financier à une banque tout en conservant son risque de longévité.

Ce premier chapitre est l'occasion de présenter brièvement le contexte des différentes parties qui constituent le corps de la thèse et de résumer les contributions des différents chapitres. La dernière partie correspondant à l'application des deux premières parties aux domaines de la démographie et de l'assurance vie.

1.1 Étude du modèle de Wishart dans le monde des actions et dans le monde des taux d'intérêt

Le modèle de Black-Scholes introduit en 1973 a permis l'évaluation de nombreux produits financiers et la couverture des risques qui leur sont associés au travers de la gestion des grecques. Les conditions de marchés se sont complexifiées et il n'est pas possible d'utiliser directement le modèle de Black-Scholes qui considère que la volatilité est constante dans le temps. Cependant, le modèle de Black-Scholes reste un modèle de référence pour le calcul de la volatilité implicite. Connaissant un prix de marché, la volatilité implicite correspond à la volatilité qu'il faudrait mettre dans le modèle de Black-Scholes afin de récupérer les prix de marchés. Contrairement à la volatilité historique qui capte l'évolution dans le passé, la volatilité implicite représente des anticipations sur les variations futures du sous-jacent et reflète le "prix du risque" associé à l'option. En pratique, la volatilité implicite des produits dérivés dépend du prix d'exercice et de la maturité de l'option : cette observation est appelée l'effet smile. Le skew est défini comme la pente de la courbe du smile en fonction du prix d'exercice dont l'évolution peut être assez complexe. Afin de capter ces phénomènes de marché, de nombreux modèles ont été mis en place avec en particulier l'apparition des

modèles à volatilité stochastique permettant de tenir compte de la dynamique du smile dans le temps.

Parmi les nombreux modèles à volatilité stochastique, il existe la classe des modèles affines qui a été introduite dans la finance par Duffie et al. [1] et dont l'utilisation pour l'évaluation d'options se fait généralement via l'étude de la transformée de Fourier [3]. Dans ces modèles, ce sont les facteurs du modèle qui ont une certaine dynamique tandis que les paramètres sont constants dans le temps. La calibration est donc assez robuste puisqu'il n'y a pas de changement de paramètres d'un jour à l'autre mais plutôt une évolution des facteurs. En effet, si nous étudions un phénomène de marché comme le skew du smile et que ce dernier dépend de paramètres qui évoluent dans le temps, il n'y a pas de stratégie de couverture pour gérer ce risque qui est complètement exotique. Il faudrait donc trouver un modèle assez flexible et avec un nombre suffisant de facteurs afin de capter au mieux les phénomènes de marché.

Nous nous intéressons dans cette partie à l'étude de modèles affines dans le monde des actions et dans le monde des taux d'intérêt. Ces modèles utilisent le processus de Wishart et auront des propriétés spécifiques répondant à certaines problématiques intéressantes.

1.1.1 Étude du modèle de Wishart dans le monde des actions

Le monde des actions nécessite une modélisation assez précise de l'évolution d'un sous-jacent afin de capter les informations disponibles sur le marché. De nombreux modèles sont apparus dont les modèles à volatilité stochastique et un de ces modèles est le modèle de Heston qui a connu un succès important. Le modèle de Heston (1993), qui est très utilisé en pratique, est un modèle dans lequel la volatilité possède une composante de retour à la moyenne. Cependant, il ne permet pas de capter à la fois les données de marchés sur les courtes et les longues maturités et ce modèle manque de flexibilité puisque la dynamique du skew est assez contrainte. Il existe de nombreuses extensions pour répondre à ces problèmes et nous proposons dans ce chapitre un modèle qui pourrait apporter des solutions.

Ainsi, nous allons proposer un modèle affine qui possède des propriétés intéressantes vis à vis de certains phénomènes de marché. Le modèle que nous étudions utilise le processus de Wishart qui est de plus en plus utilisé dans le monde de la finance. Introduits par M.F Bru [4] et adapté à la finance par C. Gourieroux et al. [5, 6], les processus de Wishart possèdent de nombreuses propriétés intéressantes. Ils ont d'ailleurs conduit à l'introduction de modèles très intéressants comme le modèle de Wishart introduit par J. Da Fonseca et al. [7]. Ce modèle à volatilité stochastique possède la propriété d'avoir une corrélation stochastique relativement "flexible" ce qui permettra de décrire la dynamique de smile et en particulier la dynamique du skew qui est stochastique.

Dans ce chapitre, nous étudions ce modèle de Wishart dans le monde des actions avec un rappel de ses nombreuses propriétés dont la propriété de corrélation stochastique. Ensuite, nous décrivons une méthode d'évaluation des options vanilles par la méthode de Fast Fourier Transform (FFT) puisque la transformée de Fourier est facilement calculable dans les modèles affines. Nous proposons aussi une méthode d'évaluation de produits plus exotiques par Monte Carlo en utilisant le lien entre le processus de Wishart et des processus d'Ornstein-Uhlenbeck.

Cependant, dans une perspective de calibration, nous cherchons à trouver des approximations en formules fermées afin de générer le smile rapidement. De nombreux travaux existent sur les formules d'approximations de smile avec la perturbation des PDE, l'approximation en courtes maturités, les projections géométriques, l'utilisation du calcul de Malliavin et encore d'autres. Nous avons donc mis en place une méthode d'approximations de smile qui s'inspire des méthodes de perturbations singulières introduites en finance par J.P. Fouque [8, 9, 10]. Nous avons adapté ces méthodes aux modèles affines en effectuant un développement de la transformée de Fourier. Ces méthodes de perturbations permettent de perturber le processus de Wishart de manière intéressante de sorte que sa loi stationnaire soit non dégénérée. Cette méthode fournit des approximations de smile en maturités courtes et en maturités intermédiaires avec deux échelles de volatilité. Les approximations en formules fermées se sont révélées efficaces et très précises et ont aussi permis l'étude de l'influence des différents paramètres du modèle. En effet, les modèles affines possèdent des paramètres constants dans le temps et il apparaît nécessaire d'avoir une connaissance précise de l'impact de chacun des paramètres.

Tout d'abord, nous analysons l'impact des coefficients diagonaux des paramètres qui ont un impact majeur sur le smile (niveau de smile, pente, convexité). Ensuite, nous testons l'influence des paramètres non diagonaux spécifiques à la structure wishart. Nous observons des effets très intéressants comme des effets sur les queues de smiles (aux petits prix d'exercices et aux grands prix d'exercices).

1.1.2 Étude du modèle de Wishart dans le monde des taux d'intérêt

Le monde des taux d'intérêt est un domaine très complexe étant donné l'aspect multidimensionnel de la modélisation. En théorie, il faudrait tenir compte de l'ensemble des taux de marché qui forment la courbe de taux : il y a un très grand nombre de taux avec les taux courts termes et les taux longs termes et il faut modéliser les corrélations des taux entre eux. Cela pose des problèmes majeurs aux praticiens qui doivent trouver un moyen de modéliser efficacement la dynamique de la courbe des taux. Les modèles les plus classiques sont le modèle de Vasicek qui modélise le taux court par un processus de Cox Ingersoll Ross

et le modèle de Heath Jarrow Merton (HJM) [12] qui diffuse la courbe des taux entièrement.

Ces modèles sont assez simples à mettre en place mais ne permettent pas de retrouver l'ensemble des prix de marchés. De nombreux modèles ont vu le jour permettant de prendre en compte les données de marchés comme les modèles de type quadratique gaussien, les modèles à volatilité stochastiques dont le plus standard est le SABR introduit par Hagan et al. [11]. L'appellation du modèle provient des paramètres Sigma, Alpha, Beta et Rho qui le caractérisent. Le SABR modélise la volatilité par une distribution log normale ce qui explique la tension des marchés : le fait de considérer des queues de distributions épaisses permet aux banques de pousser leur profit et le modèle SABR apparaît être un modèle conservateur que l'on peut tester explicitement.

Bien qu'il ait été démontré que les modèles affines ne pourront jamais atteindre des queues de distribution comme celles du SABR, ils permettent de modéliser le smile au travers de ses facteurs et apportent des solutions à de nombreux problèmes de modélisation. En ajoutant des facteurs, les modèles affines ont la capacité intrinsèque d'intégrer des sources de risques et d'identifier l'impact des différents facteurs. Dans les modèles classiques, les paramètres dépendent du temps afin pouvoir être calibrés sur de nombreux produits. Dans le cadre affine, les paramètres ne dépendent pas du temps ce qui permet de conserver une propriété de stationnarité. Le processus de calibration apparaît plus cohérent : le modèle n'est plus calibré tous les jours puisque les variations sont dues aux variations des facteurs et non aux variations de paramètres. Ces modèles sont de plus très avantageux numériquement puisque le cadre affine fournit des calculs quasi explicites au travers de résolution de systèmes de Riccati.

Cependant, les modèles affines sont complexes et sont donc assez peu utilisés en pratique. De nombreuses études sont menées afin de faciliter leur utilisation. Nous introduisons un modèle affine du taux court qui utilise le processus de Wishart. Ce modèle se révèle être une extension multidimensionnelle du modèle de Heston et permet de capter des effets de corrélations. Nous développons des méthodes d'évaluation en formules quasi fermées pour les caplets et les swaptions qui sont des produits classiques dans le monde des taux d'intérêt. Ces expressions nécessitent tout de même des méthodes numériques de type Runge Kutta puisque les coefficients du système de Riccati dépendent du temps. Ensuite, nous nous intéressons à l'influence des paramètres du modèle sur les données de marché comme la surface de volatilité implicite des swaptions à la monnaie (ATM). Enfin, nous proposons une adaptation de la méthode de perturbations régulières à ce cadre complexe permettant d'avoir une approximation de prix de caplets.

1.2 Risque de longévité et dynamique de population

Depuis près d'un siècle, nous assistons à allongement spectaculaire de l'espérance de vie des individus dans le monde entier résultant d'une baisse importante du niveau de mortalité. Cette augmentation de la durée moyenne de la vie humaine a de nombreuses causes comme l'amélioration des conditions de vie, les progrès médicaux et technologiques, les changements de mœurs et d'alimentation. Quelles sont les conséquences de ce changement majeur dans le monde ?

La conséquence principale au niveau démographique est le vieillissement de la population mondiale, de manière encore plus significative dans les pays riches. En plus d'un vieillissement par le haut avec une augmentation de l'espérance de vie, il y a aussi un vieillissement par le bas avec une baisse importante de la natalité. Ainsi, la situation heureuse d'une amélioration de la longévité des individus a aussi des conséquences importantes à l'échelle mondiale.

En plus du caractère mondial et important de cet allongement de la durée de vie, un phénomène majeur concerne l'incertitude liée à la baisse de mortalité. En effet, de nombreux démographes et scientifiques essaient de comprendre la manière dont la mortalité des individus est susceptible d'évoluer dans le futur. Il apparaît très difficile de décrire précisément son évolution puisqu'elle est dite stochastique avec des variations assez irrégulières. L'incertitude sur l'évolution de la mortalité est appelée *risque de mortalité* et l'incertitude liée à l'évolution de la longévité est appelée *risque de longévité*. Nous observons que ces risques sont corrélés l'un à l'autre puisque la longévité d'un individu est fortement liée à son taux de mortalité dans les années à venir. Cependant, il existe des différences majeures entre ces deux types de risques.

Le risque de mortalité est un risque d'évolution variationnelle dans la mesure où la mortalité varie d'une année à l'autre de manière aléatoire. Certaines variations brutales peuvent être liées à des événements importants comme une canicule ou autre. La canicule qui a créé une surmortalité importante pendant une année n'a pas pour autant contribué à une baisse de l'espérance de vie.

Le risque de longévité est un risque de tendance long terme qui lisse les variations annuelles et qui s'intéresse à une évolution de l'espérance de vie. De plus, il faut modéliser l'évolution jointe d'un groupe d'individus ce qui nécessite un modèle de dynamique de populations.

Dans cette partie de la thèse, nous nous intéressons principalement à la modélisation du risque de mortalité et du risque de longévité. Nous rappelons dans un premier temps les différents enjeux liés à la mortalité et à la longévité dans un rapport d'expertise. Nous proposons ensuite une modélisation d'évolution de la mortalité au niveau individuel qui tient

compte d'un ensemble d'informations sur l'individu. Enfin, nous introduisons un modèle mathématique de dynamique de population qui aborde le problème de l'évolution d'une population ou d'un groupe d'individus.

1.2.1 Introduction à la problématique sur la longévité (Survey)

Avec le vieillissement de la population, la longévité des individus apparaît comme un risque de plus en plus important. En effet, l'allongement de la durée de vie a des conséquences majeures au niveau social, financier, économique, politique, démographique pour n'en citer que quelques uns. Il apparaît nécessaire de mieux comprendre ce risque, de le modéliser et de trouver des moyens de le gérer et de le transférer. Dans cette partie, nous allons essayer de rappeler les notions assez standards sur le risque de mortalité et de longévité.

Tout d'abord, nous rappelons des notions standards sur l'observation de ces risques au travers des tables de mortalité et des tables de longévité. À partir des tables de mortalité nationale publiées chaque année, il est possible de capter des informations sur le niveau de mortalité et sur la longévité des individus. Il est assez difficile de constituer des tables par manque de données (surtout aux âges élevés) et qu'il faut mettre en place des méthodes de lissage pour fermer les tables.

Ensuite, nous abordons la modélisation du risque de longévité en rappelant de nombreux modèles de projections de mortalité. En effet, à partir de ces tables de mortalité statiques, il est possible de mettre en place des tables de mortalité prospectives qui tiennent compte de la tendance de l'évolution ainsi que des variations de mortalité dans les années précédentes. Les modèles les plus classiques sont les modèles de Lee Carter [13] et les modèle de type Cairns [2] qui possèdent des propriétés intéressantes de projections. De nombreuses pistes d'améliorations existent comme le fait de modéliser la mortalité par cause de décès (causes médicales par exemple) ou de tenir compte des caractéristiques individuelles comme le niveau socioéconomique, le niveau de santé,...

Après avoir rappelé les propriétés et la modélisation du risque de longévité, il faut aussi aborder la gestion de ce risque et les procédures de régulation. Les institutions financières telles que les fonds de pension et les gouvernements nationaux doivent faire face à ce risque de longévité. La situation actuelle nous invite à considérer la longévité comme un problème majeur qui a un impact sur l'ensemble des institutions financières, sur l'économie, sur la croissance. De nouvelles régulations, comme le projet européen Solvabilité 2 (Solvency 2), sont mises en place afin de contrôler ce risque et faire en sorte que les institutions financières ne soient pas trop exposées à ce risque. Nous détaillons ces nouvelles normes de régulations et la manière dont les compagnies doivent gérer ce risque.

Enfin, il faut aborder le problème du transfert de risque de longévité. De nombreuses institutions comme les assurances possèdent un risque financier important lié à leur détention de produits d'assurance vie. La plupart de ces institutions cherchent à transférer ce risque afin de limiter leur exposition à ce risque. Nous rappelons les principales méthodes de transfert de risque qui mettent en jeu la création d'indices de longévité, l'existence de produits sur le marché de la longévité et la titrisation. Ce marché émergent se met en place petit à petit et est susceptible de prendre une place très importante sur les marchés financiers. Enfin, nous rappelons les problèmes de modélisation liés à l'évolution des produits qui tiennent compte de la longévité. C'est un domaine assez complexe de recherche et nous rappelons les principaux problèmes de modélisation ainsi que les solutions proposées actuellement.

1.2.2 Modèle de mortalité par âge et par traits

Ces dernières années, la mortalité des individus a fortement baissé sur l'ensemble des tranches d'âge. Ce phénomène global possède de nombreuses causes en particulier les progrès médicaux et l'amélioration des conditions de vie. Nous observons une évolution tendancielle de la baisse de mortalité avec la présence de ruptures de tendances dans le temps et avec des périodes pour lesquelles la tendance est relativement constante. De plus, il existe des variations d'une année à l'autre dont il faut capter la structure de volatilité.

Dans cette partie, nous abordons la modélisation de type Cairns de la mortalité en détaillant un modèle standard de mortalité par âge. Nous étudions le principe de ce modèle, nous le calibrons et nous effectuons des tests de "backtesting" afin de trouver la plage de données "optimale" utilisée pour la calibration. Ce modèle est assez intéressant dans la mesure où il tient compte à la fois d'une amélioration de la longévité indépendante de l'âge et d'une amélioration de la longévité proportionnelle à l'âge.

Nous cherchons à étendre ce type de modèle en considérant un taux de mortalité dépendant de l'âge et d'autres caractéristiques. Nous rappelons tout d'abord quelles sont les variables individuelles et environnementales explicatives de mortalité. Nous nous basons sur des études économiques et démographiques qui établissent le lien entre ces variables et la mortalité.

Ensuite, nous définissons le modèle de mortalité par âge et par caractéristiques. Il existe d'autres facteurs déterminants de mortalité autres que l'âge comme les facteurs socioéconomiques. Nous calibrons le modèle sur les tables de mortalité nationale (comme pour le modèle de Cairns classique) et sur des données spécifiques qui tiennent compte de nombreuses caractéristiques. Ces données spécifiques proviennent de l'Échantillon Démographique Permanent de l'INSEE et fournissent des informations intéressantes sur la mortalité par âge et

par traits.

Ayant calibré le modèle, il est possible de faire des projections de mortalité dans le futur. Nous avons étudié l'impact de certaines caractéristiques disponibles comme le statut matrimonial, le niveau d'éducation et la catégorie socioprofessionnelle. Nous analysons l'impact de ces caractéristiques sur la mortalité des hommes et des femmes et nous étudions aussi l'évolution de cet impact dans le temps. En effet, nous savons que certaines caractéristiques ont un impact mais nous essayons aussi de capter la tendance d'évolution de cette influence qui varie avec les changements de comportements et de mœurs.

1.2.3 Modèle microscopique de dynamique de population

La modélisation de l'évolution de monde vivant et des populations est un domaine de recherche assez complexe. Il faut prendre en compte de nombreux événements comme les naissances, les morts et les migrations qui sont difficiles à modéliser et à prédire. Il existe deux grandes catégories de modèles : les modèles macroscopiques et des modèles microscopiques.

Les modèles macroscopiques étudient la dynamique d'une population à l'échelle de la population alors que les modèles microscopiques s'intéressent à une description de la population à l'échelle de l'individu. Ces modélisations abordent le sujet de deux manières différentes et ont des avantages et des inconvénients l'un par rapport à l'autre. Nous avons fait le choix d'une modélisation microscopique.

En considérant des travaux mathématiques sur la dynamique de population, nous avons essayé de modéliser l'évolution de la population d'êtres humains dans le temps. Nous présentons un modèle microscopique dans lequel chaque individu est modélisé par son âge et ses traits (qui sont des variables individuelles caractéristiques). Nous caractérisons des taux démographiques individuels (éventuellement aléatoires) qui vont dépendre des caractéristiques de l'individu. De plus, nous prenons en compte l'évolution des traits des individus dans le temps.

Nous définissons dans cette partie le processus d'évolution de la population à valeurs dans l'espace des mesures ponctuelles. Nous introduisons le formalisme nécessaire à la description du modèle mathématique qui possède de nombreuses propriétés intéressantes. Nous pouvons prouver l'existence et l'unicité forte d'un tel processus. Nous nous intéressons aussi à la méthode de simulation du processus qui permet de générer des scénarios d'évolution. De plus, nous prouvons l'existence de moments, nous caractérisons son générateur infinitésimal, et nous exprimons un problème martingale vérifié par le processus d'évolution. Ces propriétés sont d'une grande utilité et permettent de décrire la dynamique d'évolution de

la population.

Tout d'abord, l'existence d'une densité de la distribution de la population par âge et par traits assure un caractère assez régulier du processus. De plus, il est possible d'établir un lien micro/macro pour les grandes populations. En effet, lorsque la taille de la population tend vers l'infini, le modèle microscopique converge vers un modèle macroscopique et il est possible de caractériser la densité de la distribution de la population de manière explicite. En effet, nous prouvons que cette dernière suit une équation aux dérivées partielles stochastiques qui généralise l'équation aux dérivées partielles de McKendrick et Von-Foerster [14, 15] qui est classique en démographie. Cela assure une certaine cohérence du modèle à la fois à l'échelle microscopique mais aussi à l'échelle macroscopique.

1.3 Applications à la démographie et aux produits d'assurance vie

La problématique concernant le risque de longévité est très étendue avec des applications dans des domaines différents. Nous allons considérer deux applications des modèles proposés dans la partie précédente qui sont d'une part les projections démographiques avec l'étude du problème des retraites et d'autre part l'étude du risque financier de longévité ainsi que les possibilités de transfert de risque.

La baisse importante de la mortalité ainsi que le vieillissement de la population a des conséquences démographiques importantes. La population est amenée à vieillir inéluctablement et il apparaît intéressant de pouvoir quantifier ce vieillissement dans le futur en décrivant des scénarios d'évolution. L'application naturelle de ce modèle est l'étude du problème des retraites avec l'analyse de l'état de la situation et des différentes solutions possibles. Nous étudions les solutions d'une immigration choisie et d'une réforme de l'âge de la retraite du point de vue démographique.

Les institutions financières, les assurances et les fonds de pensions sont fortement exposés au risque de longévité. Les portefeuilles de rentes et les produits d'assurance vie assurent à l'assuré des flux jusqu'à son décès. Il est donc naturel que la baisse de mortalité constitue un risque majeur pour les compagnies d'assurance qu'il faut évaluer. Les compagnies d'assurance sont aussi exposées à un risque d'une autre nature appelé au risque de base qui provient du fait que les assurés vivent généralement plus longtemps que la population nationale moyenne. Nous analysons le risque de base et nous proposons un moyen de le quantifier.

Enfin, les produits d'assurance vie possèdent un risque lié à la longévité des individus et intègrent aussi un risque de taux d'intérêt qui leur est inhérent. En effet, l'échéancier des

flux de rentes s'étale sur un horizon assez long et la valorisation des produits des rentes dépend fortement du niveau des taux d'intérêt. L'assurance est prête à gérer le risque de longévité mais aimerait transférer ce risque de taux d'intérêt. Nous verrons qu'il est possible de mettre en place un produit financier de transfert de risque entre une compagnie d'assurance et une institution bancaire qui pourrait satisfaire les deux contreparties.

1.3.1 Prévisions démographiques et études du problème des retraites

En associant le modèle microscopique de dynamique de populations présenté au chapitre 6 au modèle de mortalité individuelle du chapitre 5, nous pouvons effectuer des projections démographiques de la population française. Nous avons choisi une modélisation microscopique pour les nombreux avantages qu'elle comporte. En effet, la modélisation microscopique considère un échantillon (éventuellement fictif) de la population et modélise chacun des individus avec ses caractéristiques. La modélisation microscopique fournit un panel de scénarios possibles d'évolution dans le futur. Nous verrons que le scénario moyen obtenu par cette modélisation est encore plus pessimiste que les prévisions de l'INSEE concernant le vieillissement de la population.

La situation démographique actuelle laisse entrevoir une situation difficile pour le paiement des retraites dans le futur. En France, le régime des retraites est un système par répartition et solidarité intergénérationnelle. Par conséquent, le rapport des actifs sur les inactifs, appelé "*potential support ratio*" ou "*dependency ratio*", constitue une information capitale pour le paiement des retraites. Nous étudions les possibilités proposées avec une politique d'immigration et une réforme sur l'âge de départ des retraites.

Nous cherchons à voir quelle stratégie d'immigration le gouvernement peut choisir afin d'avoir une situation démographique acceptable pour le paiement des retraites. Nous verrons que la politique d'immigration actuelle n'est pas du tout suffisante et qu'il faudrait une immigration en masse afin de modifier la structure de la population. Cependant, les chiffres étant assez irréalistes, l'idée que l'immigration puisse à elle seule régler le problème des retraites est une illusion.

Nous avons considéré la deuxième solution d'une réforme politique sur l'âge de départ à la retraite que la plupart des gouvernements occidentaux essaient de mettre en place actuellement. En effet, repousser l'âge de départ à la retraite permettrait dans l'idéal de garder les gens en activité plus longtemps. Les tests tendent à prouver qu'une réforme sur l'âge de la retraite associée à une immigration assez importante mais raisonnable (légèrement supérieure à celle des années 1970) pourrait permettre à la population d'avoir une situation démographique "acceptable" jusqu'en 2050. Ces tests ne sont pas donnés dans le but d'être appliqués tels quels mais plutôt afin de donner des pistes de réflexion sur le problème

des retraites. En effet, le problème des retraites ne se limite pas à la notion de "potential support ratio" et il faudrait prendre en compte des notions économiques comme le taux de chômage et la productivité.

1.3.2 Applications à la finance : étude du risque de base d'un portefeuille et structuration d'un produit

Les produits d'assurance vie contiennent un nombre assez important de risques financiers. Le risque principal est un risque de taux d'intérêt couplé à un risque de longévité mais il existe des produits compliqués qui contiennent aussi un risque de marché sous-jacent comme les "*variable annuities*". Nous nous intéressons aux portefeuilles de rentes qui sont les produits d'assurance vie les plus classiques.

La valorisation d'un portefeuille de rentes peut être assez complexe car elle tient compte à la fois des taux d'intérêt et de la longévité des assurés. Il faut par conséquent une modélisation du risque de longévité assez fine ainsi qu'une modélisation des taux d'intérêt qui soit satisfaisante à long terme (les horizons pouvant être assez longs). Nous utilisons le modèle de mortalité du chapitre 5 et le modèle de dynamique de population du chapitre 6. Il se pose un autre problème concernant le risque spécifique de longévité qui est l'évaluation du risque de base. En effet, les assurés ayant des caractéristiques différentes de celles d'un individu moyen ont un niveau de mortalité différent. Les assurances doivent mettre un capital pour faire face à ce risque qui est appelé risque de base. En utilisant les modèles de mortalité et de dynamique de population, nous pouvons quantifier le risque de base en tenant compte de l'hétérogénéité du portefeuille.

Le transfert de risque associé aux produits d'assurance vie est assez complexe étant donné les caractéristiques spécifiques du risque de longévité. Il est assez difficile de mettre en place un transfert de risque de longévité entre deux contreparties du fait de l'absence d'indice de référence sur le marché de la longévité qui engendre une dissymétrie d'information.

Afin de mettre en place un produit entre deux institutions, il faut que chaque contrepartie s'engage dans un domaine où elle a les compétences suffisantes. L'assurance est spécialiste dans la gestion du risque de longévité mais ne couvre que très rarement les risques de taux d'intérêt. La banque est elle spécialisée dans la gestion du risque financier comme le risque de taux d'intérêt mais ne peut supporter un risque de longévité trop important. Nous mettons en place un produit, appelé la "Longevity Nominal Chooser Swaption" (LNCS), permettant à l'assurance de transférer le risque de taux d'intérêt du produit d'assurance vie tout en conservant le risque de longévité. La Longevity Nominal Chooser Swaption est un produit qui répond aux attentes principales des deux contreparties. Nous décrivons l'étude de ce produit en détails.

Bibliographie

- [1] Duffie D., D. Filipovic and W. Schachermayer (2003), “Affine processes and applications in finance”, *Annals of Applied Probability*, 13, 984-1053.
- [2] Blake D., A. Cairns and K. Dowd (2006), “A two-factor model for stochastic mortality with parameter uncertainty : Theory and calibration”, *Journal of Risk and Insurance*, 73, 687-718
- [3] Carr P. and D. Madan (1999), “Option valuation using the fast Fourier transform”, *Journal of Computational Finance*, 2, 61-73.
- [4] Bru M.F. (1991), “Wishart Processes”, *Journal of Theoretical Probability*, 4, 724-743.
- [5] Gouriéroux C. (2006), “Wishart Processes for Stochastic Risk”, *Econometric Reviews, special issue on Stochastic Volatility*, 25, 1-41.
- [6] Gouriéroux C. and R. Sufana (2007), “Derivative Pricing with Multivariate Stochastic Volatility : Application to Credit Risk”, *Working Paper CREF 04-09*, <http://www.yorku.ca/rsufana/dpmsv.pdf>.
- [7] Da Fonseca J., M. Grasselli and C. Tebaldi (2005), “Wishart Multi-dimensional stochastic volatility”, *Working paper ESILV RR-31*.
- [8] Fouque J.P., G. Papanicolaou, R. Sircar and K. Solna (2003), “Singular Perturbations in Option Pricing”, *SIAM Journal on Applied Mathematics*, 63, 5, 1648-1665.
- [9] Fouque J.P., G. Papanicolaou, R. Sircar and K. Solna (2003), “Multiscale stochastic volatility asymptotics”, *SIAM Journal Multiscale Modeling and Simulation*, 2, 1, 22-42.
- [10] Fouque J.P. and C.H. Han (2005), “Evaluation of Compound Options using Perturbation Approximation”, *Journal of Computational Finance*, 9, 1.
- [11] Hagan P., D. Kumar, A.S. Lesniewski and D. Woodward (2002), “Managing Smile Risk”, *Wilmott Magazine*, 84-108.
- [12] Heath D., A. Jarrow and A. Morton (1992), “Bond Pricing and the Term Structure of Interest Rates : A New Methodology for Contingent Claims Valuation”, *Econometrica*, 60, 1, 77-105.
- [13] Lee R.D. and L.R. Carter (1992), “Modeling and forecasting U.S. mortality”, *Journal of the American Statistical Association*, 87, 659-675.

BIBLIOGRAPHIE

- [14] McKendrick A.G (1926), “Applications of mathematics to medical problems”, *Proc. Edin. Math.Soc.*, Vol 54, 98-130.
- [15] Von Foerster H., (1959), “Some remarks on changing populations”, *The Kinetics of Cellular Proliferation*, In Grune & Stratton, editor, New York, 382-407.

Chapitre 2

Wishart Stochastic Volatility

Introduction

This chapter is inspired by an article [4] submitted at Journal of Computational Finance that deals with the study of a Wishart stochastic volatility models. We present the framework of the model and approximations of the smile in a slow volatility time scale and a multiscale volatility.

Motivations

The Black-Scholes model (1973) does not incorporate the observable phenomenon that implied volatility of derivative products is strike- and maturity- dependent. To reproduce some market conditions like the smile effect, various models have been introduced such as local volatility and stochastic volatility models. The first step was the introduction by Dupire (1994) of local volatility models where the underlying volatility $\sigma(t, S)$ depended on the level of the underlying S itself. The most famous one is the Constant Elasticity of Variance model (CEV) in which the volatility is proportional to S^α , where α is a positive constant. Then, stochastic volatility models appeared where volatility is assumed to be a stochastic process. Thus, models became more complex because of the market incompleteness which implies that traders can not hedge their products by dealing only with the asset. The Heston model (1993) presents a volatility with an effect of mean reversion, and is commonly used in financial markets because of its flexibility. However, these models cannot fit accurately market data for short or long maturities, and recent researches have been carried out to improve this point.

A way to solve this problem is the introduction of a multifactor stochastic volatility model. C. Gourieroux and R. Sufana [21, 22] (2006) developed a multifactor version of the Heston model. Indeed, they assumed that volatility follows a Wishart process introduced by M.F. Bru [5] (1991) so that the model preserves its linear properties and consequently its tractability. J. Da Fonseca, M. Grasselli and C. Tebaldi [9](2005) have improved the

initial modeling considering that the volatility of the asset is the trace of a Wishart process. That allows to take into account stochastic correlation between the underlying asset and the volatility process and provides a wealthy but complex model. The simplistic case where the matrix of mean reversion, the volatility of volatility matrix and the correlation matrix are diagonal matrices gives a small intuition of the model performance given that the diagonal components of the Wishart process are in fact Cox-Ingersoll-Ross processes : by considering the volatility as the trace of the Wishart process, the model is equivalent to a simple multifactor Heston model. However, one can see that a multifactor Heston model (like in [7]) is not flexible enough in regard to the stochastic correlation. Indeed, in a classic extension of a multifactor Heston model, the factors appearing in the stochastic correlation formula are exactly the same as the volatility expression's ones. In order to extend this model, we will focus on a specification of the Wishart volatility model allowing to add freedom degrees concerning the stochastic correlation. Besides, we will consider the case where the correlation matrix is not diagonal and we will highlight the fact that this model differs from a standard multifactor Heston model given that the stochastic correlation depends on a new factor.

This model belongs to the class of affine models and is tractable in the way that it exists an explicit expression of the conditional characteristic function which allows the use of the Fast Fourier Transform method (FFT) given by P. Carr and D. Madan [6] (1999). Being factorial models, affine models have consequently intrinsic capabilities to integrate source of risk and are powerful tools of projection. Besides, with its mean-reversion effect, this Wishart volatility model is compatible with classic macroscopic phenomena.

Nevertheless, generating the smile of implied volatility is not fast enough, therefore the calibration remains a problem. Accurate approximations are needed in order to calibrate directly or to furnish a relevant initial set of parameters for initiating the calibration procedure. Moreover, for the users of this model such as traders, having instantly an approximation with a closed formula for the smile could be very useful. Indeed, these explicit expressions give an intuition of the influence of the parameters like correlation, volatility of volatility and mean-reversion parameters. Many academics and practitioners have tackled this challenge and have introduced specific methods. To cite some of them, there are the singular and regular perturbations methods based on the perturbation of the evaluation PDE [16, 17, 18], a procedure considering the link between the spot and the implied volatility for a long volatility scale [15], a method based on geometric using the heat kernel expansion on a Riemann manifold [29], a method based on Malliavin calculus [28] and so on. These methods have not been studied and done for the Wishart volatility model but the standard singular perturbations methods, allowing an approximation for a two-scale volatility, becomes quickly hard to apply (the second order is really difficult to obtain) [3].

Presentation of the study

This paper is organized as follows : we first present the definition and the framework of the Wishart volatility model where the volatility of the asset is given by the trace of a Wishart process. Then, we will also see some numerical ideas to evaluate options in the Wishart volatility model by Monte Carlo methods. Indeed, without simulating directly the general Wishart process, a change of probability based on a link with Ornstein-Uhlenbeck processes allows an approximation by importance function. In Section 3, we will study the asymptotic behaviour of the smile for a slow (or long) volatility time scale and for a two-scale volatility by considering the "characteristic life times" of the volatility as detailed in the article. Inspired by the singular and regular perturbations methods developed by Fouque and al. described in [16, 17, 18] to the Wishart volatility model, we have found a useful asymptotic approximation. This method uses the affine properties of the model through a perturbation of the Riccati equations and is easy to carry on. Those simple expressions allow to obtain instantly the skew and the smile for some reasonable range of parameters. Moreover, the convergence and the robustness in relation to parameters of this approximation are analyzed in practice but not in theory. Finally, this approximation provides an adequate calibration tool for admissible parameters as well as an outline of the parameters' influence in the Wishart volatility model.

Notations

To begin with, let us introduce some notations on sets of matrices :

- $\mathcal{M}_{n,m}(\mathbb{R})$: set of real matrices $n \times m$.
- $GL_n(\mathbb{R})$: set of real invertible matrices $n \times n$.
- $O_n(\mathbb{R})$: set of real orthogonal matrices $n \times n$.
- $\sigma(M)$: set of the eigenvalues of the matrix M , called spectrum of the matrix M .
- $S_n(\mathbb{R})$: set of real symmetric matrices.
- $S_n^-(\mathbb{R})$: set of nonpositive symmetric matrices i.e. for $M \in S_n^-(\mathbb{R})$, $\forall \lambda \in \sigma(M)$, $\lambda \leq 0$.
- $S_n^+(\mathbb{R})$: set of nonnegative symmetric matrices i.e. for $M \in S_n^+(\mathbb{R})$, $\forall \lambda \in \sigma(M)$, $\lambda \geq 0$.
- $\tilde{S}_n^+(\mathbb{R})$: set of strictly nonnegative symmetric matrices i.e. for $M \in \tilde{S}_n^+(\mathbb{R})$, $\forall \lambda \in \sigma(M)$, $\lambda > 0$.
- A^\top is the transposed matrix of A .
- For A a nonnegative symmetric matrix, \sqrt{A} is the unique nonnegative symmetric matrix such as $\sqrt{A}\sqrt{A} = A$.
- For $A \in \mathcal{M}_{n,n}(\mathbb{R})$, $Tr(A)$ is the trace of the matrix A .

Let us consider a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ equipped with a filtration $\{\mathcal{F}_t\}$ satisfying usual conditions. The probability measure \mathbb{Q} corresponds to a "risk-neutral" measure under which

the price of any asset is the conditional expectation of its discounted future cash flows.

2.1 The Wishart Process

The Wishart process which was originally studied by Bru [5] in 1991 was introduced in finance by Gouriéroux and al. in [21]. Then, many authors have developed stochastic volatility models using the Wishart processes and this paper is focused on the "Wishart volatility model" presented by Da Fonseca and al. [9, 11]. Some specifications and properties of the Wishart process are recalled for a better understanding.

2.1.1 Definition of the Wishart processes

Definition

The Wishart standard distribution is a multidimensional generalization of the χ^2 distribution and is very useful for the estimation of the covariance matrices in multivariate statistics [30]. Indeed, the Wishart distribution corresponds to the law of the covariance matrix estimator of a Gaussian vector sample : Let X_1, \dots, X_n be n independent Gaussian vectors in \mathbb{R}^p such as $X_i \sim \mathcal{N}(0, \Sigma)$, $i = 1, \dots, n$.

The law of the $p \times p$ random matrix $S = \sum_{i=1}^n X_i X_i^\top$ is called *Wishart distribution* noted $S \sim W(\Sigma, p, n)$. When $X_i \sim \mathcal{N}(\mu, \Sigma)$, $i = 1, \dots, n$, the law of $S = \sum_{i=1}^n X_i X_i^\top$ is called *non central Wishart distribution* noted $S \sim W(\Sigma, p, n, \mu)$.

In order to model covariance matrices dynamics, we need to focus on process valued on nonnegative matrices. Then, the trace of such matrices may be considered for modelling positive stochastic volatility process. Consider now the classic definition of the Wishart process through its diffusion equation.

Definition 2.1 (Bru [5]). Let $\{W_t, t \geq 0\}$ denote a $n \times n$ matrix-valued Brownian motion under the probability measure \mathbb{Q} . The matrix-valued process V is said to be a Wishart process if it satisfies the following diffusion equation

$$dV_t = (\beta Q^\top Q + MV_t + V_t M^\top)dt + \sqrt{V_t}dW_t Q + Q^\top dW_t^\top \sqrt{V_t}, \quad V_0 = v_0, \quad (2.1)$$

where $Q \in GL_n(\mathbb{R})$ is a $n \times n$ invertible matrix, M is a $n \times n$ nonpositive matrix, $v_0 \in \tilde{S}_n^+$ is a strictly nonnegative symmetric matrix and β a real such as $\beta > (n - 1)$.

The condition $\beta > n - 1$ is introduced to ensure existence and uniqueness of the solution $V_t \in S_n^+$ of Equation (2.1). Moreover, as shown in [5, 9], for $\beta \geq n + 1$, eigenvalues of the solution are strictly nonnegative $\forall t \geq 0$ a.s $V_t \in \tilde{S}_n^+$

The Laplace transform

The diffusion satisfied by the Wishart process belongs to the class of affine diffusions [13]. Consequently, an explicit expression for the Laplace transform is available via the resolution of a Riccati system and this notion will be a key tool for future developments. This Laplace transform belongs to the family of Wishart distribution Laplace transforms as described for example in [30] what explains the designation "Wishart process". Let $\Psi_t(h, \Theta) = \mathbb{E}[\exp(-\text{Tr}(\Theta V_{t+h})) | V_t]$ define the Laplace transform of the Wishart process.

Proposition 2.1 (Gourieroux & Sufana [21]). *For $\Theta \in S_n^+$ and $t, h \geq 0$, let $\Psi_t(h, \Theta) = \mathbb{E}[\exp(-\text{Tr}(\Theta V_{t+h})) | V_t]$ the Laplace transform of V_{t+h} given V_t . Then,*

$$\Psi_t(h, \Theta) = \frac{\exp(-\text{Tr}[\Delta(h)^\top \Theta (I_n + 2\Sigma(h)\Theta)^{-1} \Delta(h) V_t])}{(\det[I_n + 2\Sigma(h)\Theta])^{\frac{\beta}{2}}}, \quad (2.2)$$

with

$$\Delta(h) = \exp(hM), \quad \Sigma(h) = \int_0^h \Delta(s) Q^\top Q \Delta(s)^\top ds.$$

2.1.2 Examples

Let us present some useful and simple examples.

Multidimensional Cox-Ingersoll-Ross processes

Let us consider a Wishart process with diagonal matrices M , Q and R . It can be seen as a multidimensional Cox-Ingersoll-Ross process. Indeed, the dynamics of the diagonal components are :

$$dV_t^{ii} = (\beta(Q^{ii})^2 + 2M^{ii}V_t^{ii})dt + 2Q^{ii} \sum_{k=1}^n \sqrt{V_t^{ki}} dW_t^{ki}.$$

Although their dynamics seem to depend on other components of the Wishart process, that is just an illusion. The Brownian motions can be concatenated through the following expression

$$d \langle V^{ii}, V^{ii} \rangle_t = 4 \left((Q^{ii})^2 \left(\sum_{k=1}^n \sqrt{V_t^{ki}} \sqrt{V_t^{ki}} \right) \right) dt = 4(Q^{ii})^2 (\sqrt{V_t} \sqrt{V_t})^{ii} dt = 4(Q^{ii})^2 V_t^{ii} dt.$$

So, $B = (B^1, \dots, B^n)$ defined by $dB_t^i = \sqrt{V_t^{ii}}^{-1} \sum_{k=1}^n (\sqrt{V_t^{ki}}) dW_t^{ki}$ is a vector of n independent Brownian motions. Then,

$$dV_t^{ii} = (\beta(Q^{ii})^2 + 2M^{ii}V_t^{ii}) dt + 2Q^{ii} \sqrt{V_t^{ii}} dB_t^i.$$

Then, in this simple case where parameters are diagonal matrices, the diagonal components of the Wishart process are independent Cox-Ingersoll-Ross processes.

Link between Wishart processes and Ornstein Uhlenbeck processes

A natural way to generate a "Wishart process" is to replace the Gaussian vectors X_i in the definition of a Wishart distribution (Section 1.1.1) with Ornstein-Uhlenbeck processes $X_{i,t}$. With this representation which implies that β is an integer, the Wishart volatility model described in Section 2 appears as an example of Quadratic Gaussian models. Considering β independent n -dimensional Ornstein-Uhlenbeck processes :

$$dX_{k,t} = MX_{k,t}dt + Q^\top dW_{k,t}, \quad \text{and} \quad V_t := \sum_{k=1}^{\beta} X_{k,t}X_{k,t}^\top$$

Then, the process (V_t) is a Wishart process with the following dynamics.

$$dV_t = (\beta Q^\top Q + MV_t + V_t M^\top)dt + \sqrt{V_t}dW_t Q + Q^\top dW_t^\top \sqrt{V_t},$$

with W a matrix-valued Brownian motion, determined by

$$\sqrt{V_t}dW_t = \sum_{k=1}^{\beta} X_{k,t}dW_{k,t}^\top.$$

In this framework, the matrix M can be seen as the mean-reversion parameter of the Wishart process and Q as the volatility parameter.

Consequently, the simulation Ornstein-Uhlenbeck processes allows the simulation of a Wishart process in the case where β is an integer. Therefore, the first step is to consider the simple case where $\beta \geq n + 1$ is an integer. For the initialization of the Ornstein-Uhlenbeck processes, one can notice that the initial matrix V_0 is real and symmetric and can be then diagonalized. Using the eigenvectors Φ_k and the eigenvalues λ_k , one can write :

$$V_0 = \sum_{k=1}^n \lambda_k \Phi_k \Phi_k^\top.$$

Then, the initialization of the Ornstein-Uhlenbeck processes can be done

$$\begin{cases} \bar{X}_{k,0} &= \sqrt{\lambda_k} \Phi_k \text{ if } 1 \leq k \leq n, \\ \bar{X}_{k,0} &= 0 \text{ if } n+1 \leq k \leq \beta \end{cases}$$

Finally, for $1 \leq i \leq N$, simulating $\{\bar{X}_{k,t_i}, 1 \leq k \leq \beta\}$ by one of the methods described in the next paragraph furnishes the discretization of this Wishart process

$$\bar{V}_{t_i} = \sum_{k=1}^{\beta} \bar{X}_{k,t_i} \bar{X}_{k,t_i}^\top.$$

This simulation procedure for this Wishart process needs the simulation of Ornstein-Uhlenbeck processes which is classic and can be done by an Euler scheme or an exact

simulation using the knowledge of the conditional distribution (See Appendix A).

However, this way does not provide the most general class of the Wishart process and there is a very useful extension in practice by considering the same dynamics with a real β . As developed in Section 1.3.2, Girsanov theorem gives a change of probability measure allowing to move from β integer to β real. The general case will be detailed after.

2.1.3 Determinant dynamics and change of probability measure

In this part, the stochastic differential equation describing the determinant dynamics of Wishart process is presented and analyzed. At the end of this section, a change of probability measure is highlighted allowing a change of drift in the dynamics of Wishart process. Using determinant dynamics, the *likelihood ratio martingale* has an explicit functional form (2.4) and allows an efficient evaluation procedure by importance function.

Determinant dynamics

The dynamics of determinant will be essential to handle this change of probability measure that will be very important for the evaluation of expectations where β is not an integer. A useful theorem given in [5] has to be recalled (without proof) for the future understanding.

Theorem 2.1 (Bru [5]). *Let $\{\zeta_t = (\zeta_t^{ij})_{1 \leq i, j \leq n}\}$ be a continuous semimartingale process valued in the matrix set $S_n(\mathbb{R})$. Let $U_t \in O_n(\mathbb{R})$ such as $U_t^\top \zeta_t U_t = \text{diag}(\lambda_t^1, \dots, \lambda_t^n)$ where $\{\lambda_t^i, 1 \leq i \leq n\}$ are the eigenvalues of ζ_t .*

Let $\{A_t = (A_t^{ij})_{1 \leq i, j \leq n}, t \geq 0\}$ be the process defined as

$$A_t = \int_0^t U_s^\top d\zeta_s U_s, \forall t \geq 0 \quad \text{such as} \quad d \langle A^{ij}, A^{ji} \rangle = \Gamma_t^{ij} dt.$$

Then, the dynamics of eigenvalues of V_t can be written as

$$d\lambda_t^i = dM_t^i + dJ_t^i, \quad dJ_t^i = \sum_{j \neq i} \frac{1}{\lambda_t^i - \lambda_t^j} \Gamma_t^{ij} dt + d\Theta_t^i,$$

where dM^i and $d\Theta^i$ are respectively the martingale part and the finite variation part of dA^{ii} .

Let us denote $\tilde{Q}_t = U_t^\top Q U_t$ and $\tilde{M}_t = U_t^\top M U_t$ and $Z_t = \int_0^t U_s^\top dW_s U_s$.

By applying the previous theorem, the dynamics of eigenvalues are characterized by the following stochastic differential equation.

$$d\lambda_t^i = [\beta(\tilde{Q}_t^\top \tilde{Q}_t)^{ii} + 2\lambda_t^i \tilde{M}_t^{ii} + \sum_{j \neq i} \frac{(\lambda_t^i \tilde{Q}_t^\top \tilde{Q}_t)^{jj} + (\lambda_t^j \tilde{Q}_t^\top \tilde{Q}_t)^{ii}}{\lambda_t^i - \lambda_t^j}] dt + 2\sqrt{\lambda_t^i} \sum_{k=1}^n \tilde{Q}_t^{ki} dZ_t^{ik}.$$

It can be proved that all eigenvalues are independent and are non-colliding [5, 14] meaning that $d \langle \lambda^i, \lambda^j \rangle_t = 0$. Therefore, the differentiation of the determinant yields to the

following expression $d(\det(V_t)) = \det(V_t) \sum_{i=1}^n \frac{d\lambda_t^i}{\lambda_t^i}$. Moreover, by using the trace invariance by a change of basis, the dynamics of determinant of V_t are obtained by

$$\frac{d(\det(V_t))}{\det(V_t)} = [(\beta - n + 1)Tr(V_t^{-1}Q^\top Q) + 2Tr(M)]dt + 2Tr(\sqrt{V_t^{-1}}dW_tQ),$$

which implies that the

$$d \log(\det(V_t)) = [(\beta - n - 1)Tr(V_t^{-1}Q^\top Q) + 2Tr(M)]dt + 2Tr(\sqrt{V_t^{-1}}dW_tQ). \quad (2.3)$$

Change of the probability measure

From a mathematical point of view, it is really important to analyze the change of the probability measure allowing a change of the drift in the process dynamics. In practice, for a financial use like calibration of parameters on historic data, a Wishart process has to be studied in its general form with $\beta \geq n + 1$ and $\beta \in \mathbb{R}$. Considering the floor function allows to write $\beta = K + 2\nu$ with $K = [\beta] \geq n + 1$ and ν a real number such as $0 \leq \nu \leq \frac{1}{2}$.

The aim of this development is to find a change of probability measure in order to change the generalized Wishart diffusion into the simple one where K is an integer. This new probability measure is noted $\mathbb{Q}^*_{|\mathcal{F}_T}$ and can be expressed as follows

Theorem 2.2. *Let $q = K + \nu - n - 1$. If $f_T(\mathbb{Q}, \mathbb{Q}^*) = \frac{d\mathbb{Q}_{|\mathcal{F}_T}}{d\mathbb{Q}^*_{|\mathcal{F}_T}}$ defines the Radon-Nikodym derivative of $\mathbb{Q}_{|\mathcal{F}_T}$ with respect to $\mathbb{Q}^*_{|\mathcal{F}_T}$, then*

$$f_T(\mathbb{Q}, \mathbb{Q}^*) = \frac{\det(V_T)^{\frac{\nu}{2}}}{\det(V_0)^{\frac{\nu}{2}}} \exp[-\nu T Tr(M)] \exp \left[-\frac{\nu}{2} q \int_0^T Tr(V_s^{-1}Q^\top Q) ds \right]. \quad (2.4)$$

Proof 2.1. Like in [14], a probability measure \mathbb{Q}^* can be specified through an exponential martingale

$$\frac{d\mathbb{Q}}{d\mathbb{Q}^*_{|\mathcal{F}_T}} = \exp \left[\nu \int_0^T Tr(\sqrt{V_s^{-1}}dW_sQ) - \frac{\nu^2}{2} \int_0^T Tr(V_s^{-1}Q^\top Q) ds \right].$$

Therefore, this expression suggests to define a new process W^* by

$$W_t^* = W_t + \nu \int_0^t \sqrt{V_s^{-1}}Q^\top ds.$$

By Girsanov theorem, it is easy to check that W^* is a matrix-valued Brownian motion under the probability measure \mathbb{Q}^* . Consequently, the dynamics of Wishart process under this probability measure \mathbb{Q}^* can be rewritten as follows

$$dV_t = (KQ^\top Q + MV_t + V_tM^\top)dt + \sqrt{V_t}dW_t^*Q + Q^\top(dW_t^*)^\top \sqrt{V_t}.$$

The Radon-Nikodym derivative can be simplified using the determinant dynamics. Indeed, as noted in the first section, we have

$$\log\left[\frac{\det(V_t)}{\det(V_0)}\right] = 2T \operatorname{Tr}(M) + (K - n - 1) \int_0^T \operatorname{Tr}(V_s^{-1} Q^\top Q) ds + 2 \int_0^T \operatorname{Tr}(\sqrt{V_s^{-1}} dW_s^* Q).$$

Finally, the change of the probability measure can be obtained by

$$d\mathbb{Q}|_{\mathcal{F}_T} = \frac{\det(V_T)^{\frac{\nu}{2}}}{\det(V_0)^{\frac{\nu}{2}}} \exp[-\nu T \operatorname{Tr}(M)] \exp\left[-\frac{\nu}{2}(K + \nu - n - 1) \int_0^T \operatorname{Tr}(V_s^{-1} Q^\top Q) ds\right] . d\mathbb{Q}|_{\mathcal{F}_T}^*,$$

which completes the proof. \square

This change of probability measure allows to bring back the problem to the simple case considered before. It will be very useful for many applications and more particularly for evaluation by Monte Carlo method.

2.2 Wishart stochastic volatility model in the stock derivatives market

2.2.1 Presentation of the Wishart volatility model

In this part, the Wishart volatility model is presented as well as its essential characteristics. The properties inherent in this model will allow an accurate and an efficient consistency with the market. The underlying volatility is defined as the trace of a Wishart process. Therefore, the diagonal components of the Wishart matrix will be the factors guiding the dynamics of volatility. Under the risk neutral probability measure, the asset dynamics are given by the following expression

$$\begin{aligned} \frac{dS_t}{S_t} &= rdt + \operatorname{Tr}[\sqrt{V_t}(dW_t R + dZ_t \sqrt{I_n - R R^\top})] & S_0 &= x. \\ dV_t &= (\beta Q^\top Q + M V_t + V_t M^\top)dt + \sqrt{V_t} dW_t Q + Q^\top dW_t^\top \sqrt{V_t} & V_0 &= v_0, \end{aligned} \quad (2.5)$$

where $\{V_t, t \geq 0\}$ is a Wishart matrix-valued process as introduced in the previous section, r is the interest rate considered constant, S_t the price of the asset at date t , R a matrix such as $\rho(R) = \max\{\lambda, \lambda \in \sigma(R)\} \leq 1$, W , Z are independent matrix-valued Brownian motions.

The matrix R describes to the correlation between the Brownian matrices of the asset and those of the Wishart process. The matrix M represents the matrix of mean reversion and the matrix Q is the volatility of the volatility.

The dynamics of log-price $Y_t = \log(S_t)$ are deduced easily

$$\begin{aligned} dY_t &= \left(r - \frac{\operatorname{Tr}(V_t)}{2}\right)dt + \operatorname{Tr}[\sqrt{V_t}(dW_t R + dZ_t \sqrt{I_n - R R^\top})] & Y_0 &= y. \\ dV_t &= (\beta Q^\top Q + M V_t + V_t M^\top)dt + \sqrt{V_t} dW_t Q + Q^\top dW_t^\top \sqrt{V_t} & V_0 &= v_0. \end{aligned} \quad (2.6)$$

It is conventional to write the log-price and its volatility dynamics as driven by two correlated Brownian motions B_s^Y and B_s^V . In fact, as it is done in [11] by using a concatenation of the Brownian motions like in Section 1.2, the asset dynamics can be rewritten as follows

$$\begin{aligned} dY_t &= \left(r - \frac{\text{Tr}(V_t)}{2}\right)dt + \sqrt{\text{Tr}(V_t)}dB_t^Y, \\ d\text{Tr}(V_t) &= (\beta\text{Tr}(Q^\top Q) + 2\text{Tr}(MV_t))dt + 2\sqrt{\text{Tr}(Q^\top QV_t)}(\rho_t dB_t^Y + \sqrt{1 - \rho_t^2}dB_t^V), \end{aligned} \quad (2.7)$$

where $d \langle B^Y, B^V \rangle_t = \rho_t = \frac{\text{Tr}(R^\top QV_t)}{\sqrt{\text{Tr}(V_t)}\sqrt{\text{Tr}(Q^\top QV_t)}}$.

Fourier transform

In order to calculate price of options by FFT method, define the instantaneous realized mean variance $U_t = \int_0^t \text{Tr}(V_s)ds$, and consider the Laplace Transform $\Psi_t(T, \theta, \gamma, \delta) = \mathbb{E}[e^{T\text{Tr}(\theta V_t) + \gamma Y_T + \delta U_T} | \mathcal{F}_t]$ of (V, Y, U) .

This function has to be explicitly calculated in order to price many derivative products such as vanilla options or Variance Swap. Since the system (V, Y, U) defines an affine model, the Laplace Transform Ψ can be easily expressed with functions that are solutions of a Riccati system [13]. The resolution is possible for example by using the linearization method introduced by J. Da Fonseca and al.[9].

Theorem 2.3.

$$\Psi_t(T, \theta, \gamma, \delta) = \exp[\text{Tr}[A(T-t)V_t] + B(T-t)Y_t + C(T-t)U_t + D(T-t)],$$

where A, B, C and D are solutions of the following Riccati equations

$$\begin{aligned} A'(\tau) &= \left(\frac{\gamma(\gamma-1)}{2} - \delta\right)I_n + A(\tau)M + (M^\top + 2\gamma R^\top Q)A(\tau) + 2A(\tau)Q^\top QA(\tau). \\ B(\tau) &= \gamma, \quad C(\tau) = \delta. \\ D'(\tau) &= r\gamma + \beta\text{Tr}[Q^\top QA(\tau)], \end{aligned} \quad (2.8)$$

with the terminal conditions $A(0) = \theta$ and $D(0) = 0$.

By using the linearization method, the solutions for A and D are given by

$$\begin{aligned} A(\tau) &= F(\tau)^{-1}G(\tau). \\ D(\tau) &= \gamma\tau[r - \beta\text{Tr}(R^\top Q)] - \frac{\beta}{2}[\text{Tr}(M)\tau + \log(\det F(\tau))], \end{aligned} \quad (2.9)$$

with

$$\begin{aligned} [G(\tau) \ F(\tau)] &= [0 \ I_n] \exp(\tau Z(\gamma, \delta)) \\ Z(\gamma, \delta) &= \begin{pmatrix} M & -2Q^\top Q \\ \left(\frac{\gamma(\gamma-1)}{2} - \delta\right)I_n & -(M^\top + 2\gamma R^\top Q) \end{pmatrix}. \end{aligned}$$

Proof 2.2. The function $f_t(T, V, Y, U) = \exp[Tr[A(T-t)V_t] + B(T-t)Y_t + C(T-t)U_t + D(T-t)]$ has to be a martingale. Thus, Girsanov theorem suggests to find a martingale M_t such as

$$\log[f_t(T, V, Y, U)] = M_t - \frac{1}{2} \langle M_t, M_t \rangle. \quad (2.10)$$

Applying Itô formula to the affine function $f_t(T, V, Y, U)$ yields

$$\begin{aligned} d \log[f_t(T, V, Y, U)] = & - \left(Tr[A'(T-t)V_t]dt + B'(T-t)Y_t + C'(T-t)U_t + D(T-t) \right) dt \\ & + \left(Tr[\beta Q^\top Q A(T-t)] + Tr[A(T-t)M V_t] + Tr[M^\top A(T-t)V_t] \right) dt \\ & + Tr[A(T-t)\sqrt{V_t}dW_t Q] + Tr[A(T-t)Q^\top dW_t^\top \sqrt{V_t}] \\ & + \left(B(T-t) - \frac{1}{2}Tr[B(T-t)V_t] \right) dt \\ & + Tr[B(T-t)\sqrt{V_t}dW_t R] + Tr[B(T-t)dZ_t \sqrt{I_n - RR^\top}] \\ & + C(T-t)Tr(V_t)dt. \end{aligned} \quad (2.11)$$

This expression allows us to find the martingale M_t

$$\begin{aligned} dM_t = & Tr[A(T-t)\sqrt{V_t}dW_t Q] + Tr[A(T-t)Q^\top dW_t^\top \sqrt{V_t}] \\ & + Tr[B(T-t)\sqrt{V_t}dW_t R] + Tr[B(T-t)dZ_t \sqrt{I_n - RR^\top}]. \end{aligned}$$

For the calculation of the quadratic variation of M_t , noted $\langle M_t, M_t \rangle$, let us recall some useful expressions

$$\begin{aligned} d \langle Tr(AdW_t B) \rangle &= Tr(BAA^\top B^\top)dt, \\ d \langle Tr(AdW_t B), Tr(CdW_t D) \rangle &= Tr(DCA^\top B^\top)dt. \end{aligned}$$

Then, the quadratic variation $\langle M_t, M_t \rangle$ of the martingale is expressed as follows

$$\begin{aligned} d \langle M_t, M_t \rangle = & \left(4Tr[A(T-t)Q^\top Q A(T-t)V_t] + B^2(T-t)Tr[V_t] \right) dt \\ & + \left(Tr[B(T-t)R^\top Q A(T-t)V_t] \right) dt. \end{aligned}$$

Then, replacing this expression in (2.10) and identifying with (2.11) gives

$$\begin{aligned} 0 = & \left(-2Tr[A(T-t)Q^\top Q A(T-t)V_t] - \frac{1}{2}B^2(T-t)Tr[V_t] - \frac{1}{2}Tr[B(T-t)R^\top Q A(T-t)V_t] \right) \\ & - \left(Tr[A'(T-t)V_t]dt + B'(T-t)Y_t + C'(T-t)U_t + D(T-t) \right) \\ & + \left(Tr[\beta Q^\top Q A(T-t)] + Tr[A(T-t)M V_t] + Tr[M^\top A(T-t)V_t] \right) \\ & + \left(B(T-t) - \frac{1}{2}Tr[B(T-t)V_t] \right) + C(T-t)Tr(V_t). \end{aligned}$$

Finally, identifying the terms with V_t , Y_t and U_t gives

$$\begin{aligned} A'(\tau) &= \left(\frac{\gamma(\gamma-1)}{2} - \delta \right) I_n + A(\tau)M + (M^\top + 2\gamma R^\top Q)A(\tau) + 2A(\tau)Q^\top QA(\tau) \\ B(\tau) &= \gamma, \quad C(\tau) = \delta \\ D'(\tau) &= r\gamma + \beta \text{Tr}[Q^\top QA(\tau)]. \end{aligned}$$

Riccati equation linearization

Let us search A such as

$$A(\tau) = (F(\tau))^{-1}G(\tau), \quad F(0) = I_n, \quad G(0) = 0.$$

By replacing this expression in the first Riccati equation of (2.12)

$$\begin{aligned} -(F(\tau))^{-1}F'(\tau)A(\tau) + (F(\tau))^{-1}G'(\tau) &= \left(\frac{\gamma(\gamma-1)}{2} - \delta \right) I_n + (F(\tau))^{-1}G(\tau)M, \\ &+ (M^\top + 2\gamma R^\top Q)A(\tau) \\ &+ 2(F(\tau))^{-1}G'(\tau)Q^\top QA(\tau). \end{aligned}$$

Therefore, by multiplying on the left by $F(\tau)$ and identifying terms with A and without A , we obtain

$$\begin{aligned} G'(\tau) &= G(\tau)M + \left(\frac{\gamma(\gamma-1)}{2} - \delta \right) F(\tau), \\ F'(\tau) &= -G(\tau)Q^\top Q - F(\tau)(M^\top + 2\gamma R^\top Q). \end{aligned} \tag{2.12}$$

This system of equations can be written as follows

$$\begin{bmatrix} G'(\tau) & F'(\tau) \end{bmatrix} = \begin{bmatrix} G(\tau) & F(\tau) \end{bmatrix} Z(\gamma, \delta),$$

with $Z(\gamma, \delta) \in \mathcal{M}_{2n}(\mathbb{R})$ defined by

$$Z(\gamma, \delta) = \begin{pmatrix} M & -2Q^\top Q \\ \left(\frac{\gamma(\gamma-1)}{2} - \delta \right) I_n & -(M^\top + 2\gamma R^\top Q) \end{pmatrix}.$$

Thus, the solution is given by

$$\begin{bmatrix} G(\tau) & F(\tau) \end{bmatrix} = \begin{bmatrix} 0 & I_n \end{bmatrix} \exp(\tau Z(\gamma, \delta)).$$

Resolution of the equation that D follows

By integrating the fourth equation of (2.12), we have

$$D(\tau) = \gamma r\tau + \beta \text{Tr} \left[\int_0^\tau Q^\top QA(s) ds \right].$$

On the other hand, by multiplying on the left left by $F(\tau)^{-1}$ the second equation of (2.12), integrating it between 0 and τ , and taking the trace, we find

$$\beta Tr[\int_0^\tau Q^\top QA(s)ds] = -\frac{\beta}{2}(Tr[\int_0^\tau F(s)^{-1}F'(s)ds] + Tr(M^\top + 2\gamma R^\top Q)\tau).$$

Finally the solution is given by

$$D(\tau) = \gamma\tau[r - \beta Tr(R^\top Q)] - \frac{\beta}{2}[Tr(M)\tau + \log(\det F(\tau))].$$

□

In order to calculate the price of Call options with a resolution by FFT method, a closed formula for the Fourier transform of log-price is needed. The expression of this Fourier transform is defined by $\Phi_t(T, \gamma) = \mathbb{E}[\exp(i\gamma Y_T)|\mathcal{F}_t] = \Psi_t(T, 0, i\gamma, 0)$ and an explicit formula can be deduced easily. Another proof of this corollary can be found in [9]

Corollary 2.1.

$$\Phi_t(T, \gamma) = \exp[Tr[A(T-t)V_t] + B(T-t)Y_t + C(T-t)],$$

where A and C are solutions of the following Riccati equations

$$\begin{aligned} B(\tau) &= i\gamma. \\ A'(\tau) &= \frac{i\gamma(i\gamma-1)}{2}I_n + A(\tau)M + (M^\top + 2i\gamma R^\top Q)A(\tau) + 2A(\tau)Q^\top QA(\tau). \\ C'(\tau) &= ir\gamma + \beta Tr[Q^\top QA(\tau)]. \end{aligned}$$

The linearization method can be used and gives

$$\begin{aligned} A(\tau) &= F(\tau)^{-1}G(\tau), \\ C(\tau) &= i\gamma\tau[r - \beta Tr(R^\top Q)] - \frac{\beta}{2}[Tr(M)\tau + \log(\det F(\tau))], \end{aligned} \quad (2.13)$$

with

$$\begin{aligned} [G(\tau) \ F(\tau)] &= [0 \ I_n] \exp(\tau Z(\gamma)) \\ Z(\gamma) &= \begin{pmatrix} M & -2Q^\top Q \\ \frac{i\gamma(i\gamma-1)}{2}I_n & -(M^\top + 2i\gamma R^\top Q) \end{pmatrix}. \end{aligned}$$

Flexibility of the Wishart volatility model

As seen in Section 1.2.1, the Wishart process is an extension of a multidimensional Cox-Ingersoll-Ross process. Thus, the Wishart volatility model will be also an extension of a

multidimensional Heston model by considering diagonal matrices. Indeed, there is a bijection concerning the volatility parameters between the Heston model and the Wishart volatility model with this restricted specification. Moreover, we have the same correlation structure for those models and the stochastic correlation is given by

$$\rho_t = \frac{\sum_{i=1}^n R^{ii} Q^{ii} V_t^{ii}}{\sqrt{\sum_{i=1}^n V_t^{ii}} \sqrt{\sum_{i=1}^n (Q^{ii})^2 V_t^{ii}}}.$$

This particular case underlines the fact that the Wishart volatility model extends the multifactor Heston model in terms of volatility and correlation. In practice, the model used will be a simple extension of the multifactor Heston model. Moreover, the Wishart volatility model in the general case has additional properties. Indeed, in a multidimensional Heston model, the correlation is stochastic but depends on factors that generate the volatility dynamics. In the case of the Wishart volatility model, the correlation depends on the volatility factors but depends also on other factors in formula (2.8) (the non diagonal components of the Wishart process V_t). It is an important property allowing degrees of freedom for the correlation and consequently for the skew and the smile. For example, in the Heston model, the change of sign of the skew is constrained by the correlation coefficients and the volatility factors whereas in the Wishart volatility model, the change of sign does not have this constraint thanks to the additional independent factor [10].

To develop the intuition on the importance of the stochastic correlation, we refer to asymptotic results in short maturities from V. Durrleman [15] where the link between correlation and implied volatility skew at the money are explicit.

Asymptotic skew in the Wishart volatility model

The Wishart volatility model gives a stochastic correlation (for $n \geq 2$) which is an important property. The correlation ρ_t being stochastic allows the model to generate a smile with a wealthy structure. Indeed, the market is more and more complex and now the skew is considered by practitioners as stochastic. Following standard notations with $\Sigma_t(T, K)$ denotes the implied volatility of a Call with a maturity T and a strike K , the skew at the money can be calculated for short maturities by using asymptotic methods developed by Durrleman.

Theorem 2.4. (V. Durrleman [15])

Let us suppose that we are given adapted processes a, \tilde{a}, b, c such that there exists a strictly positive solution to the following stochastic differential equation :

$$d\sigma_t^2 = (b_t + a_t^2 + \sigma_t^2(a_t + c_t))dt - 2\sigma_t(a_t dW_t + \tilde{a}_t d\tilde{W}_t),$$

where a given initial condition σ_0 and such that

$$d < a, W >_t = -\sigma_t \left(\frac{3c_t}{2} + \frac{3a_t^2}{4\sigma_t^2} - \frac{\tilde{a}_t^2}{\sigma_t^2} \right) dt.$$

Further let

$$S_t = S_0 \exp \left(\int_0^t \sigma_s dW_s - \frac{1}{2} \int_0^t \sigma_s^2 ds \right)$$

be the stock process and $\Sigma_t(T, K)$ be the corresponding implied volatility, then

$$\begin{aligned} \Sigma_t(t, S_t) &= \sigma_t, & \frac{\partial \Sigma_t}{\partial K}(t, S_t) &= -\frac{a_t}{2S_t \Sigma_t(t, S_t)}, & \frac{\partial \Sigma_t}{\partial T}(t, S_t) &= \frac{b_t}{4\Sigma_t(t, S_t)}, \\ \frac{\partial^2 \Sigma_t}{\partial K^2}(t, S_t) &= \frac{1}{2S_t^2 \Sigma_t(t, S_t)} \left(a_t + c_t - \frac{a_t^2}{2\Sigma_t(t, S_t)^2} \right). \end{aligned} \quad (2.14)$$

Therefore, the skew being strongly linked with the correlation represented by $\frac{a_t}{\sigma(t)}$, this correlation has to be stochastic. Then, the need of a model that can explain this phenomenon becomes essential. The stochastic correlation appears to be one of the crucial notions describing the capacity of a model to generate smiles. As explained in the next section, in terms of correlation, the Wishart volatility model appears more flexible than a multi-dimensional Heston model. For the expression of the skew at the money ($K = S_t$) in the Wishart volatility model, the application of this theorem with $a_t = -\frac{\text{Tr}(R^\top Q V_t)}{\text{Tr}(V_t)}$ provides the following expression

$$\frac{\partial \Sigma_t}{\partial K}(t, S_t) = \frac{\text{Tr}(R^\top Q V_t)}{2S_t (\text{Tr}(V_t))^{\frac{3}{2}}}. \quad (2.15)$$

Comment 2.1. *It is important to notice that Theorem 2.4 also provides an expression of $\frac{\partial^2 \Sigma_t}{\partial K^2}(t, S_t)$ and $\frac{\partial \Sigma_t}{\partial T}(t, S_t)$ allowing to analyze the convexity and the slope at the money of the asymptotic smile, but it is not brought up in this development. We focus here on the skew at the money to bring to the light the stochastic correlation property and to give a benchmark for the approximation in the next section but we could have also described the convexity and the slope at the money of the implied volatility.*

2.2.2 Evaluation of options by Monte Carlo methods

There are many products for which the evaluation does not have any closed-form expression. Therefore, a Monte Carlo method has to be considered in order to price those contracts and the problem to simulate a Wishart process becomes essential.

The use of a standard Euler scheme for a general Wishart process is prohibited since it does not guarantee the positiveness of the simulated matrices. The other method consisting in taking the positive part of the eigenvalues would introduce a bias and is rejected. Then, an efficient method, based on a fundamental link between Wishart and Ornstein-Uhlenbeck

processes, was highlighted in the previous section in the case of β real. This method will naturally keep the positiveness of the matrix-valued Wishart process [19, 1]¹. Now, the question is to see how the complete Wishart volatility model can be simulated and how the European options can be evaluated by Monte Carlo. Considering a general Wishart process, the underlying dynamics are expressed under the probability measure \mathbb{Q}^* in order to use the simulation method developed in Section 1.2.2

$$\begin{aligned}\frac{dS_t}{S_t} &= [r - \nu \text{Tr}(Q^\top R)]dt + \text{Tr}[\sqrt{V_t}(dW_t^* R + dZ_t \sqrt{I_n - RR^\top})]. \\ dV_t &= (KQ^\top Q + MV_t + V_t M^\top)dt + \sqrt{V_t}dW_t^* Q + Q^\top (dW_t^*)^\top \sqrt{V_t}.\end{aligned}$$

Integrating this stochastic differential equation furnishes a useful form of the log-price discretization

$$\begin{aligned}\log(S_{t_i}) &= \log(S_{t_{i-1}}) + \left((r - \nu \text{Tr}(Q^\top R))\Delta t - \frac{1}{2} \int_{t_{i-1}}^{t_i} \text{Tr}(V_s)ds \right) \\ &\quad + \text{Tr} \left[\int_{t_{i-1}}^{t_i} \sqrt{V_s} dW_s R + \int_{t_{i-1}}^{t_i} \sqrt{V_s} dZ_s \sqrt{I_n - RR^\top} \right].\end{aligned}$$

The increments $\{\varepsilon_{k,i}\}_{1 \leq k \leq \beta}$, already used to simulate the Ornstein Uhlenbeck processes with the scheme (2.25), allows to simulate $\text{Tr}[\int_{t_{i-1}}^{t_i} \sqrt{V_s} dW_s R]$. For the other terms, a standard Euler scheme is selected giving the following expressions

$$\begin{aligned}\int_{t_{i-1}}^{t_i} \sqrt{V_s} dW_s &= \sum_{k=1}^{\beta} \int_{t_{i-1}}^{t_i} X_{k,s} dW_{k,s}^\top = \sqrt{\Delta t} \sum_{k=1}^{\beta} X_{k,t_i} \varepsilon_{k,i}^\top, & \int_{t_{i-1}}^{t_i} V_s ds &= \bar{V}_{t_{i-1}} \Delta t. \\ \text{Tr} \left[\int_{t_{i-1}}^{t_i} \sqrt{V_s} dZ_s \sqrt{I_n - RR^\top} \right] &= \sqrt{\Delta t} \sqrt{\text{Tr}(V_{t_{i-1}}(I_n - RR^\top))} Z_i,\end{aligned}$$

where $Z_i \sim \mathcal{N}(0,1)$ are the components of an independent Gaussian vector and being independent with the increments $\varepsilon_{k,i}$

Finally, the Euler scheme for the asset is characterized by

$$\begin{aligned}\log(\bar{S}_{t_i}) &= \log(\bar{S}_{t_{i-1}}) + \left[r - \nu \text{Tr}(Q^\top R) - \frac{1}{2} \text{Tr}(\bar{V}_{t_{i-1}}) \right] \Delta t + \sqrt{\Delta t} \text{Tr} \left[\sum_{k=1}^{\beta} \bar{X}_{k,t_{i-1}} \varepsilon_{k,i}^\top R \right] \\ &\quad + \sqrt{\Delta t} \sqrt{\text{Tr}(V_{t_{i-1}}) - \text{Tr}(V_{t_{i-1}} R R^\top)} Z_i.\end{aligned}$$

Therefore, the price of an European option at the maturity T with a payoff f is evaluated by

1. During the revision August 2009 - January 2010, two papers have been posted about the simulation of the Wishart model : on September 2009 by P. Gauthier and D. Possamai on June 2010 by A. Abdelkoddousse and A. Alfonsi

$$\begin{aligned}
 \pi_0 &= \mathbb{E}_{\mathbb{Q}}[\exp(-rT)f(S_T)] \\
 &= \exp[-(r + \nu \operatorname{Tr}(M))T]\mathbb{E}_{\mathbb{Q}^*} \left\{ \left(\frac{\det(V_T)}{\det(V_0)} \right)^{\frac{\nu}{2}} \exp[-\frac{\nu}{2}(K + \nu + n - 1) \int_0^T \operatorname{Tr}(V_s^{-1} Q^\top Q) ds] f(S_T) \right\}.
 \end{aligned}$$

The simulation of the integral appearing in the Radon-Nikodym function can be discretized and approximated as follows

$$\int_t^{t+\Delta t} \operatorname{Tr}(V_s^{-1} Q^\top Q) ds \sim \frac{1}{2} \Delta t \operatorname{Tr}[(V_t^{-1} + V_{t+\Delta t}^{-1}) Q^\top Q].$$

Finally, it is possible to make a Monte Carlo evaluation for an option given that the asset and the Radon-Nikodym derivative of \mathbb{Q} with respect to \mathbb{Q}^* can be simulated using an Euler scheme. One can see that the expectation under the new probability measure depends on the path of $(V_t)_{t \leq T}$ what makes the evaluation costly.

2.2.3 Evaluation of options by FFT

The Wishart volatility model belongs to the class of affine models for which there exists a closed formula for the Fourier transform [13]. This transform, noted $\Phi_t(T, \gamma) = \mathbb{E}[\exp(i\gamma Y_T) | \mathcal{F}_t]$, $\gamma \in \mathbb{C}$, is widely used to find options prices by calculating the inverse Fourier transform by FFT.

Expression of the Call price

The premium at date t of a Call option with strike $K = \log(k)$ and maturity T is given by

$$c_t(T, K) = \exp[-r(T-t)] \mathbb{E}[(\exp(Y_T) - \exp(k))^+ | \mathcal{F}_t].$$

As it has been done in [6], the calculation of the premium is deduced by the expression of the Fourier transform. Let α define a constant such as $x \rightarrow \exp(\alpha x)(\zeta - \exp(x))^+ \in L^1(\mathbb{R})$ and in practice, Carr & Madan consider that $\alpha = 1.1$ is an empirical good value for the Heston model. It appears really important to have a function in $L^1(\mathbb{R})$ in order to use the inverse Fourier transform as explained by S. Levendorskii in [26]. Therefore, a modified price is defined as follows

$$c_t^\alpha(T, k) = \exp(\alpha k) c_t(T, k).$$

By applying a Fubini integration theorem, the Fourier transform of the modified price is given by

$$\begin{aligned}
 \Psi_t^\alpha(T, \nu) &= \int_{-\infty}^{\infty} \exp(i\nu k) c_t^\alpha(T, k) dk \\
 &= \exp[-r(T-t)] \mathbb{E} \left[\int_{-\infty}^{Y_T} \exp[(i\nu + \alpha)k] (\exp(Y_T) - \exp(k)) dk | \mathcal{F}_t \right] \\
 &= \exp[-r(T-t)] \frac{\Phi_t(T, \nu - (1 + \alpha)i)}{(1 + \alpha + i\nu)(\alpha + i\nu)}.
 \end{aligned} \tag{2.16}$$

At last, the premium can be obtained by inversion of the Fourier transform using the fact that the function $\Psi_t^\alpha(T, \nu)$ has an odd imaginary part and an even real part

$$\begin{aligned} c_t(T, k) &= \frac{\exp(-\alpha k)}{2\pi} \int_{-\infty}^{\infty} \exp(-i\nu k) \Psi_t^\alpha(T, \nu) d\nu \\ &= \exp[-r(T-t)] \frac{\exp(-\alpha k)}{\pi} \mathcal{R}e \int_0^{\infty} \exp(-i\nu k) \frac{\Phi_t(T, \nu - (1+\alpha)i)}{(1+\alpha+i\nu)(\alpha+i\nu)} d\nu. \end{aligned} \quad (2.17)$$

Numerical upgrade

In this paragraph, we consider a practical point of view and we present a variant of Formula (2.17) in order to improve numerical results. The idea is to use this formula for a simple model (like the Black-Scholes model with a constant volatility $\bar{\sigma}$) and to apply this expression for numerical applications. This improves the effectiveness of numerical integration. It is done by analogy with the variance reduction method for Monte Carlo options pricing through a control variate method using the Black-Scholes price $c_t^{BS}(T, K)$ and its Fourier transform $\Phi_t^{BS}(T, \gamma)$ calculated by :

$$\Phi_t^{BS}(T, \gamma) = \exp[i\gamma Y_t + i\gamma r(T-t) + \frac{i\gamma(i\gamma-1)}{2} \bar{\sigma}^2(T-t)]. \quad (2.18)$$

The Call price can be calculated as a perturbation of the Call price in the Black-Scholes model with a volatility $\bar{\sigma}$

$$c_t(T, k) = c_t^{BS}(T, k) + I_t(r, \bar{\sigma}, \alpha, T),$$

where $I_t(r, \bar{\sigma}, \alpha, T)$ is given by

$$I_t(r, \bar{\sigma}, \alpha, T) = \exp[-r(T-t)] \frac{\exp(-\alpha k)}{\pi} \mathcal{R}e \left[\int_0^{\infty} \exp(-i\nu k) \frac{(\Phi_t(T, \nu - (1+\alpha)i) - \Phi_t^{BS}(T, \nu - (1+\alpha)i))}{(1+\alpha+i\nu)(\alpha+i\nu)} d\nu \right].$$

We consider that a good choice for the volatility $\bar{\sigma}$ is the realized mean volatility in the Wishart volatility model

$$\bar{\sigma} = \sqrt{\frac{1}{T} \int_0^T \mathbb{E}[Tr(V_t)] dt}.$$

This volatility can be calculated as follows. Let $f(t) = \mathbb{E}(V_t)$, f is the solution of the following differential equation

$$f'(t) = \beta Q^\top Q + Mf(t) + f(t)M^\top, \quad f(0) = V_0.$$

Then, the solution can be found explicitly

$$f(t) = \exp(tM) V_0 \exp(tM^\top) + \beta \int_0^t \exp(sM) Q^\top Q \exp(sM^\top) ds.$$

Assume now that M is a symmetric matrix and take the trace

$$Tr[f(t)] = Tr[V_0 \exp(2tM)] + \beta Tr[Q^\top Q(2M)^{-1}(\exp(2tM) - I_n)].$$

Finally, by integrating between 0 and T , the realized mean variance is given by

$$\bar{\sigma}^2 = \frac{1}{T} Tr[(V_0 + \beta Q^\top Q(2M)^{-1})(2M)^{-1}(\exp(2TM) - I_n)] - \beta Tr[Q^\top Q(2M)^{-1}].$$

Comment 2.2. For a non positive matrix M , we can take the same formula with $M = \frac{M+M^\top}{2}$ which gives better results in practice.

2.3 Smile dynamics in the Wishart volatility model

The hedge of exotic options not only needs the trade of the asset but requires also the trade of options such as Call and Put options. Therefore, these vanilla options are hedging instruments and the calibration procedure has to take into account of their price dynamics. Then, models considered and used by traders have to be flexible enough to explain market prices. In other words, the model must be able to generate the observable smile effect (strike dependency of the Black-Scholes implied volatility). The level of the skew, corresponding to the derivative of the implied volatility with respect to the strike, reflects the correlation between the asset and its volatility.

Heston model is a single-factor stochastic volatility model that can be calibrated accurately on a market implied volatility surface. However, when the calibration is done on both short and long maturities smile curves, a single-factor model does not provide a flexible modelling of the volatility term structure. Indeed, a single factor model cannot take into account the variability of the skew, also known as correlation risk, and cannot explain independent fluctuations of the smirk in the level and the slope over time. A classic solution is to increase the number of factors in order to achieve this goal like in [7]. For instance, P. Christoffersen and al. [7] introduce a bidimensional extension to solve this problem. The Wishart volatility model is a multidimensional extension of the Heston model allowing an adequate correlation structure in order to reproduce the skew effect. An analysis of these features is proposed in this section.

The covariance dynamics is represented by a matrix-valued process of size n . The case $n = 2$ is enough for this study allowing to deal with the case of a two-scale volatility (medium maturities). It is also a framework where the computation is not so laborious and explicit formulas are available. In this section, the study of the Wishart volatility model and the effect of some parameters on the smile are considered. Inspired by the regular and singular perturbations methods introduced by J.P. Fouque and al. [16, 17, 18] and by standard approximation procedures [23], we present a perturbation algorithm, based on the perturbation of Riccati equations. This method will allow to generate instantly asymptotic smiles with a good level of accuracy. The case of short maturities is a relevant

case and is important for a calibration prospect. Moreover, in the two-scale volatility case, the introduction of a slow varying factor gives a much better fit for options with longer maturities [17]

2.3.1 Specification of the Wishart volatility model

It is important to describe the admissibility of the parametrization for this model. It is important to highlight a canonical model from which all the equivalent models can be obtained by invariant transformation [12]. Considering the case where $M \in S_n(\mathbb{R})$ and $Q \in GL_n(\mathbb{R})$, it is possible to define an equivalent model with M diagonal because the matrix-valued Brownian motion is invariant by any rotation. Indeed, there exists a matrix D diagonal and an orthogonal matrix P such as $D = P^\top M P$. Denoting $\tilde{V}_t = P^\top V_t P$, $\tilde{W}_t = P^\top W_t P$ and $\tilde{Q} = P^\top Q P$, an equivalent Wishart process can be deduced by this rotation

$$d\tilde{V}_t = (\beta \tilde{Q}^\top \tilde{Q} + D \tilde{V}_t + \tilde{V}_t D) dt + \sqrt{\tilde{V}_t} d\tilde{W}_t \tilde{Q} + \tilde{Q}^\top d\tilde{W}_t^\top \sqrt{\tilde{V}_t}.$$

Therefore, only eigenvalues of the mean-reversion matrix M matter in the specification of volatility dynamics. In dimension n , the parameters of the model can be written as follows :

$$M = \begin{pmatrix} -\lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & -\lambda_n \end{pmatrix}, \quad Q = \begin{pmatrix} q_{11} & \dots & q_{1n} \\ \vdots & \ddots & \vdots \\ q_{n1} & \dots & q_{nn} \end{pmatrix},$$

with $\forall i \in \{0, \dots, n\}$, $\lambda_i > 0$.

2.3.2 Condition for the perturbation method

The regular and singular perturbations methods, introduced by J.P Fouque and al. [16, 17, 18], consist in the perturbation of the parameters of a process such as the stationary distribution of this process is non-degenerate. The singular perturbations method will be useful to study the cases of different scales of volatility.

When Q is diagonal, the Wishart process is a multidimensional Cox Ingersoll Ross process and it converges when time goes to infinity to a Gaussian process whose covariance matrix is non-degenerate provided a link between the volatility and the drift. This link is the same as in the Vasicek model in interest rate modelling and the values $\nu_i = \frac{q_{ii}}{\sqrt{\lambda_i}}$ has to be kept constant. The values ν_i correspond to a ratio between volatility of volatility and the square root of the mean-reversion parameters.

Comment 2.3. *The scale of the time in the dynamics cannot be interpreted independently of the evolution of the process. Let $\nu_i = \frac{q_{ii}}{\sqrt{\lambda_i}}$. In the case where Q is diagonal, it is easy to*

calculate the following expression

$$\mathbb{E}(V_T^{ii}|\mathcal{F}_t) = \frac{\beta\nu_i^2}{2} + (V_t^{ii} - \frac{\beta\nu_i^2}{2})\exp[-2m_i(T-t)].$$

Consequently, $\tau_i = \frac{1}{2m_i}$ for $i \in \{1, \dots, n\}$ are mean-reverting times of V_T and $\frac{\beta\nu_i^2}{2}$ for $i \in \{1, 2\}$ correspond to the asymptotic mean of the diagonal components of V_T when T goes to infinity. In particular, if $T - t \ll \tau_i$ then $\mathbb{E}(V_T^{ii}|\mathcal{F}_t) \sim V_t^{ii}$ which is the case of small maturity T or slow mean-reversion volatility. If $\tau_i \ll T - t$, then $\mathbb{E}(V_T^{ii}|\mathcal{F}_t) \sim \frac{\beta\nu_i^2}{2}$ which corresponds to a long maturity T or a fast mean-reversion volatility.

In the general case, the stationary distribution can be analyzed through the limit of the Laplace transform of the Wishart process when time goes to infinity. The Laplace transform depends only on $\Sigma(h)$ and $\Delta(h)$ given in (2.2). Since $\lim_{h \rightarrow \infty} \Delta(h) = I_n$ because the mean reversion matrix M has negative eigenvalues, the equivalent condition for the invariance of the Laplace Transform at infinity is that $\Sigma(\infty) = \lim_{h \rightarrow \infty} \Sigma(h)$ is non-degenerate.

Comment 2.4. It is easy to check that in the case where β is an integer, the Wishart process can be generated by Ornstein Uhlenbeck processes with $V_t = \sum_{k=1}^{\beta} X_{k,t} X_{k,t}^{\top}$ and it appears that the covariance matrix of the Wishart process at time h depends only on the covariance matrix of the Ornstein Uhlenbeck processes which is $\Sigma(h)$ defined in (2.2).

Let us denote $\Omega = Q^{\top}Q = (\omega)_{ij} \big|_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, an easy calculation gives :

$$\Sigma_{ij}(h) = \int_0^h w_{ij} e^{-(\lambda_i + \lambda_j)s} ds = w_{ij} \frac{1 - e^{-(\lambda_i + \lambda_j)h}}{\lambda_i + \lambda_j}.$$

Then, it is easy to find that $\Sigma_{ij} = \Sigma_{ij}(\infty) = \frac{w_{ij}}{\lambda_i + \lambda_j}$ and the condition is that all the coefficients Σ_{ij} have to remain constant by the perturbation method. The condition can be written as

$$w_{ij} = (\lambda_i + \lambda_j)\Sigma_{ij}.$$

Considering that condition, in order to find the condition between M and Q with the Cholesky decomposition. Indeed, the fact that Ω is a strictly nonnegative symmetric matrix, the Cholesky transformation ensures that there exists a lower triangular matrix C such as $\Omega = CC^{\top}$ (there is uniqueness if the coefficients of C are positive). Then the matrix Q is deduced by a rotation of the matrix C and there exists a unique matrix $R \in O_n(\mathbb{R})$ such as $Q = CR$. Then it is sufficient to find a link between C and Ω to deduce the link between M and Q . The following expression

$$\Omega_{ij} = \sum_{k=1}^i C_{ik} C_{jk}$$

can be rewritten

$$\begin{aligned}\Omega_{ii} &= C_{i1}^2 + \dots + C_{ii}^2, \\ \Omega_{ij} &= C_{i1}C_{j1} + \dots + C_{ii}C_{ji}.\end{aligned}$$

This system is invertible and the resolution is done gradually beginning from the first coefficient Ω_{11} and looking down the lines and columns of the matrix Ω , one can find easily :

$$\begin{aligned}C_{ii} &= \left(\Omega_{ii} - \sum_{k=1}^{i-1} C_{ik}^2 \right)^{\frac{1}{2}} \\ C_{ji} &= \frac{\Omega_{ij} - \sum_{k=1}^{i-1} C_{ik}C_{jk}}{C_{ii}} \quad \text{for } i+1 \leq j \leq n.\end{aligned}$$

In dimension 2, this resolution is given by

$$\begin{aligned}C_{11} &= \sqrt{2\lambda_1\Sigma_{11}}. \\ C_{21} &= \frac{(\lambda_1 + \lambda_2)\Sigma_{21}}{C_{11}} = \frac{\lambda_1 + \lambda_2}{\sqrt{\lambda_1}} \frac{\Sigma_{21}}{\sqrt{2\Sigma_{11}}}. \\ C_{22} &= \sqrt{2\Sigma_2\lambda_2 - C_{21}^2} = \sqrt{2\Sigma_{22}\lambda_2 - \frac{(\lambda_1 + \lambda_2)^2}{\lambda_1} \frac{(\Sigma_{21})^2}{2\Sigma_{11}}}.\end{aligned}$$

Comment 2.5. The term $2\Sigma_{22}\lambda_2 - \frac{(\lambda_1+\lambda_2)^2}{\lambda_1} \frac{(\Sigma_{21})^2}{2\Sigma_{11}}$ is positive given that the determinant of matrix Ω is positive and is equal to $4m_1m_2\Sigma_{11}\Sigma_{22} \geq (m_1 + m_2)^2(\Sigma_{21})^2$.

2.3.3 Perturbation Method of the Riccati equations

In this part, a method of perturbations based on the affine properties of the model is presented. Indeed, the perturbation of the Riccati equations provides an asymptotic development of the Call price with report to ϵ corresponding to the perturbation of the parameters.

This method consists in the perturbation of the parameters of the volatility process like in [16, 17, 18] taking into account the affine properties of the model. In the case where the time scale of the volatility is long compared with the maturity option, this method is really similar from an asymptotic approximation in short time like in [15] but is a little different. Indeed, as noted in Section 3.1, during the perturbation process, there is a constraint between the drift and the volatility of the process (in the Wishart case, the coefficients ν_i remain constant). Therefore, in the following, this case will also be named as a case of "short maturities". Besides, singular perturbations methods can be used for a multiscale volatility framework as described in the next section.

Finally, this procedure is efficient and allows us to reach higher orders than standard perturbation methods like the singular perturbations method on partial differential equations

that appears really complicated after the first order [3]. The affine properties allow an efficient, easy, and general procedure. Remember the homogeneous Riccati system associated to the Wishart volatility model where for simplicity, we use the notation $\theta = i\gamma$

$$\begin{aligned} A'(\tau) &= \frac{\theta(\theta-1)}{2}I_n + A(\tau)M + (M^\top + 2\theta R^\top Q)A(\tau) + 2A(\tau)Q^\top QA(\tau). \\ C'(\tau) &= r\theta + \beta Tr[Q^\top QA(\tau)]. \end{aligned}$$

Then, a development of the function A furnishes instantly a development for C .

Let us consider a characteristic order ε in the perturbation and let us focus on a development of the solution $A(\tau)$ of the form

$$A(\tau) = \sum_{j=0}^{\infty} \varepsilon^{\frac{j}{2}} A^j(\tau). \quad (2.19)$$

Injecting this development in the perturbed Riccati equation and identifying terms in ε provides the expected approximation. In the following, this method will be applied for two important cases : the case of a slow volatility time scale and the case of the multiscale volatility.

2.3.4 Smile dynamics for a slow volatility time scale

Introduction

Observe the case of a slow volatility time scale with $(T-t) \ll \tau_i = \frac{1}{2m_i}$, meaning that the option maturity is much less than the mean-reverting times of the volatility components. In other words, the volatility V_t does not fluctuate a lot around its starting value. It is a short maturity smile because the maturity of the option is small in relation to the characteristic times of the volatility process. The matrix M can be written as follows :

$$\begin{aligned} M &= \varepsilon \tilde{M} \\ \tilde{M} &= \begin{pmatrix} -m_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & -m_n \end{pmatrix}. \end{aligned}$$

In order to satisfy the condition of a non-degenerate stationary distribution, it is easy to see (see Appendix B) that the matrix Q must be written :

$$\begin{aligned} Q &= \sqrt{\varepsilon} \tilde{Q} \\ \tilde{Q} &= \begin{pmatrix} q_{11} & \dots & q_{1n} \\ \vdots & \ddots & \vdots \\ q_{n1} & \dots & q_{nn} \end{pmatrix}. \end{aligned}$$

Consequently εm_i are small in relation to $\frac{1}{(T-t)}$. The method is detailed for the short maturities case and the procedure is presented for an approximation at order 1 i.e. in $(\sqrt{\varepsilon})$ and at order 2 i.e. in (ε) .

Development of the Riccati function A and C

By perturbing the Riccati equation, at order 2, a solution is looked for under the following development

$$A(\tau) = A^0(\tau) + \sqrt{\varepsilon}A^1(\tau) + \varepsilon A^2(\tau) + o(\varepsilon).$$

Consequently, injecting this formulation into the Riccati equations and doing calculations step by step, gives the expected functions

$$\begin{aligned} A^0(\tau) &= \frac{\theta(\theta-1)}{2}\tau^2 I_n, & A^1(\tau) &= \frac{\theta^2(\theta-1)}{2}\tau^2 R^\top \tilde{Q}, \\ A^2(\tau) &= \theta^3(\theta-1)\frac{\tau^3}{3}(R^\top \tilde{Q})^2 - \frac{\theta(\theta-1)}{2}\tau^2 \tilde{M} + \theta^2(\theta-1)^2\frac{\tau^3}{6}\tilde{Q}^\top \tilde{Q}. \end{aligned}$$

By using the link between functions C and A , the development for C is deduced

$$C(\tau) = r\tau\theta + \varepsilon C^2(\tau) + o(\varepsilon) \quad \text{with} \quad C^2(\tau) = \beta \frac{\theta(\theta-1)}{4}\tau^2 Tr(\tilde{Q}^\top \tilde{Q}).$$

Approximation of the Call price

By considering the previous expressions, it is possible to find a Taylor expansion of the Fourier transform of $Y_t = \log(S_t)$ with report to ε

$$\begin{aligned} \Phi(T, \theta) &= \exp \left[\frac{\theta(\theta-1)}{2}(T-t)Tr(V_t) + \theta \log(S_t) + r(T-t)\theta \right] * \\ &\quad \left[1 + \sqrt{\varepsilon}Tr(A^1 V_t) + \varepsilon \left(Tr(A^2 V_t) + C^2 + \frac{1}{2}(Tr(A^1 V_t))^2 \right) + o(\varepsilon) \right], \end{aligned}$$

with the following formula

$$\begin{aligned} Tr(A^1 V_t) &= \frac{\theta^2(\theta-1)}{2}(T-t)^2 Tr(R^\top \tilde{Q} V_t), \\ Tr(A^2 V_t) &= \theta^3(\theta-1)\frac{(T-t)^3}{3}Tr((R^\top \tilde{Q})^2 V_t) - \frac{\theta(\theta-1)}{2}(T-t)^2 Tr(\tilde{M} V_t) \\ &\quad + \theta^2(\theta-1)^2\frac{(T-t)^3}{6}Tr(\tilde{Q}^\top \tilde{Q} V_t). \end{aligned}$$

We recognize the Fourier transform in the Black-Scholes model define by (Formula 2.18) with a volatility $\bar{\sigma} = Tr(V_t)$.

This expansion at order $o(\max(\varepsilon, \delta))$ must be analyzed checking that the neglected terms have to be small in comparison to the first terms of the development. The main question of this procedure is how the price approximation can be inferred from the Fourier transform development.

In a general way, assume that there is a development for the Fourier transform under the form

$$\Phi_t(T, \gamma) = \Phi_t^{BS}(\bar{\sigma})(T, \gamma) \sum_j \varepsilon^{\frac{j}{2}} P_j(i\gamma),$$

where P_j are real polynomials. Noticing that $P[\frac{\partial}{\partial y}]\Phi_t^{BS(\bar{\sigma})}(T, \gamma) = P(i\gamma)\Phi_t^{BS(\bar{\sigma})}(T, \gamma)$ with Formula (2.18), the price of a European option is given by

$$C_t(T, k) = \sum_j \varepsilon^{\frac{j}{2}} P_j \left(\frac{\partial}{\partial y} \right) C_t^{BS(\bar{\sigma})}(T, k).$$

Moreover, the expressions of the successive derivatives of the price with respect to the log spot $y = \log(S_0)$ are needed and can be easily calculated.

Consider non standard notations with

$$a = \log(Ke^{-r(T-t)}) + \frac{1}{2}\sigma^2(T-t) \text{ what gives the link } \frac{a-y}{\sigma\sqrt{T-t}} = -d_0$$

$$\forall n \geq 2, \quad \Lambda_y^{BS,n} = \frac{\partial^n C_{BS}(\sigma)}{\partial y^n} - \frac{\partial^{n-1} C_{BS}(\sigma)}{\partial y^{n-1}}.$$

A useful recursive relation was found allowing the calculation in the Black-Scholes model of Call price successive derivatives with respect to the log spot y (see Appendix C for explicit calculation of successive derivatives)

$$\begin{aligned} \forall n \geq 4, \quad \frac{\partial^n C_{BS}}{\partial y^n} &= \frac{\partial^{n-1} C_{BS}}{\partial y^{n-1}} + \frac{(a-y)}{\sigma^2(T-t)} \left(\frac{\partial^{n-1} C_{BS}}{\partial y^{n-1}} - \frac{\partial^{n-2} C_{BS}}{\partial y^{n-2}} \right) \\ &- \frac{(n-3)}{\sigma^2(T-t)} \left(\frac{\partial^{n-2} C_{BS}}{\partial y^{n-2}} - \frac{\partial^{n-3} C_{BS}}{\partial y^{n-3}} \right). \end{aligned} \quad (2.20)$$

Using classic formula in the Black-Scholes model, the development for the Call price can be determined

$$P = P^0 + \sqrt{\varepsilon}P^1 + \varepsilon P^2 + o(\varepsilon),$$

with

$$\begin{aligned} P^0 &= C_{BS(\sigma)}, \quad \sigma = Tr(V_t). \\ P^1 &= \frac{\tau^2}{2} Tr(R^\top \tilde{Q} V_t) \Lambda_y^{BS,3}. \\ P^2 &= \beta \frac{\tau^2}{4} Tr(\tilde{Q}^\top \tilde{Q}) \Lambda_y^{BS,2} - \frac{\tau^2}{2} Tr(\tilde{M} V_t) \Lambda_y^{BS,2} + \frac{\tau^3}{6} Tr(\tilde{Q}^\top \tilde{Q} V_t) (\Lambda_y^{BS,4} - \Lambda_y^{BS,3}) \\ &+ \frac{\tau^3}{3} Tr((R^\top \tilde{Q})^2 V_t) \Lambda_y^{BS,4} + \frac{\tau^4}{8} Tr(R^\top \tilde{Q} V_t)^2 (\Lambda_y^{BS,6} - \Lambda_y^{BS,5}). \end{aligned} \quad (2.21)$$

For example, in dimension $n = 2$, an approximation of the price at order $o(\sqrt{\varepsilon})$ is highlighted

$$\begin{aligned} C_t(T, k) &= C_t^{BS}(T, k) + \sqrt{\varepsilon} \frac{(T-t)^2}{2} \left(\frac{\partial^3 C_t^{BS(\bar{\sigma})}}{\partial y^3}(T, k) - \frac{\partial^2 C_t^{BS(\bar{\sigma})}}{\partial y^2}(T, k) \right) [(R_{11}q_{11} + R_{21}q_{21})V_t^{11} \\ &+ (R_{11}q_{12} + R_{21}q_{22} + R_{12}q_{11} + R_{22}q_{21})V_t^{12} + (R_{12}q_{12} + R_{22}q_{22})V_t^{22}] + o(\sqrt{\varepsilon}). \end{aligned}$$

Approximation of the lognormal implied volatility

From a price development, an expansion of the lognormal implied volatility can be inferred. The same kind of expansion is considered

$$\Sigma = \Sigma^0 + \sqrt{\varepsilon}\Sigma^1 + \varepsilon\Sigma^2 + o(\varepsilon). \quad (2.22)$$

By injecting this development in the Black-Scholes price, a Taylor expansion around $\Sigma^{0,0}$ furnishes another expression for the Call price

$$P = C_{BS(\Sigma^0)} + \sqrt{\varepsilon}\Sigma^1 \left(\frac{\partial C_{BS(\sigma)}}{\partial \sigma} \Big|_{\sigma=\Sigma^0} \right) + \varepsilon \left[\Sigma^2 \left(\frac{\partial C_{BS(\sigma)}}{\partial \sigma} \Big|_{\sigma=\Sigma^0} \right) + \frac{1}{2} (\Sigma^1)^2 \left(\frac{\partial^2 C_{BS(\sigma)}}{\partial \sigma^2} \Big|_{\sigma=\Sigma^0} \right) \right].$$

In the following, we denote $Vega^{BS} = \frac{\partial C_{BS(\sigma)}}{\partial \sigma} \Big|_{\sigma=\Sigma^{0,0}}$ and $Vomma^{BS} = \frac{\partial^2 C_{BS(\sigma)}}{\partial \sigma^2} \Big|_{\sigma=\Sigma^{0,0}}$. Finally, by comparing this expression with the development (2.21) obtained previously and by using classic expression in the Black-Scholes model (see Appendix B), the development of the implied volatility is specified

$$\begin{aligned} \Sigma^0 &= Tr(V_0), \\ \Sigma^1 &= \frac{P^1}{Vega^{BS}}, \\ \Sigma^2 &= \frac{P^2 - \frac{1}{2} (\Sigma^1)^2 Vomma^{BS}}{Vega^{BS}}. \end{aligned}$$

Therefore, we can obtain an explicit expression of asymptotic smile at order $o(\varepsilon)$ but the formula is too long. For instance, at order $o(\sqrt{\varepsilon})$, an explicit and concise formula for the smile can be deduced

$$\begin{aligned} \hat{\Sigma}_t(T, K) &= \sqrt{Tr(V_t)} + \frac{\sqrt{\varepsilon} P^1(t, Y_t, V_t^{11}, V_t^{21}, V_t^{22})}{\frac{\partial C_t^{BS}}{\partial \sigma}(T, K, \sqrt{Tr(V_t)})} \\ &= \sqrt{Tr(V_t)} + \left[\log(K e^{-r(T-t)}) - Y_t + \frac{Tr(V_t)(T-t)}{2} \right] \\ &\quad \times \frac{[(R_{11}q_{11} + R_{21}q_{21})V_t^{11} + (R_{11}q_{12} + R_{21}q_{22} + R_{12}q_{11} + R_{22}q_{21})V_t^{12} + (R_{12}q_{12} + R_{22}q_{22})V_t^{22}]}{2(Tr(V_t))^{\frac{3}{2}}}. \end{aligned} \quad (2.23)$$

Notice that $\sqrt{Tr(V_t)}$ is the stochastic volatility and is also the implied volatility at the forward money ($K = F_t^T = e^{Y_t + r(T-t)}$) for very short maturities ($T = t$). Furthermore, an expression at order $o(\sqrt{\varepsilon})$ of the skew at the forward money is obtained and one can check that this formula corresponds to the one obtained by Durrleman for $T = t$ (2.15)

$$\frac{\partial \Sigma_t}{\partial K}(T, F_t^T) \sim \frac{Tr(R^\top Q V_t)}{2(Tr(V_t))^{\frac{3}{2}} F_t^T}.$$

This expression shows that the sign of the skew can be driven by the parameters of the Wishart model and particularly with the correlation parameters. Finally, we have obtained a development of the smile for all strikes (not just at the forward money) and in maturities. This approximation is constrained by the fact that the maturity option T has to be much less than the characteristics times of the volatility process τ_i and it is also clear that this approximation will be more accurate in the neighborhood of the forward money (Section 3.5). In the next section, we will deal with the case of options with a medium maturity.

2.3.5 Smile dynamics for a two-scale volatility

For the case of a multiscale of volatility, we have chosen the case of the dimension $n = 2$ and a matrix Q diagonal. The justification for the dimension $n = 2$ is that the calculations are really complicated. Consider two scales with parameters ε and δ and in this case, the matrix M can be written

$$M = \begin{pmatrix} -\frac{m_1}{\varepsilon} & 0 \\ 0 & -m_2\delta \end{pmatrix}.$$

A two-scale volatility is brought up what means $\tau_1 \ll (T - t) \ll \tau_2$ or in other words $\frac{1}{2m_1} \ll (T - t) \ll \frac{1}{2m_2}$. A justification for a matrix Q diagonal is given but it could exist a way to have a matrix Q full but it is not clear. Let us suppose for simplicity that $\delta = \varepsilon$ and the covariance matrix is full with $\Sigma_{21} = \Sigma_{12} \neq 0$. The expression of the Cholesky coefficient

$$\begin{aligned} C_{11} &= \sqrt{2m_1\Sigma_{11}}\sqrt{\varepsilon}, \\ C_{21} &\sim \frac{m_2}{\sqrt{m_1}} \frac{\Sigma_{21}}{\sqrt{2\Sigma_{11}}} \frac{1}{\varepsilon^{\frac{3}{2}}}, \\ (C_{22})^2 &\sim \frac{2\Sigma_{22}m_2}{\varepsilon} - \frac{(m_2)^2}{m_1} \frac{(\Sigma_{21})^2}{2\Sigma_{11}} \frac{1}{\varepsilon^3} \\ &\underset{\varepsilon \rightarrow 0}{\sim} -\frac{(m_2)^2}{m_1} \frac{(\Sigma_{21})^2}{2\Sigma_{11}} \frac{1}{\varepsilon^3} < 0, \end{aligned}$$

leads to an absurdity given that a $(C_{22})^2$ has to stay positive. Then we deduce that $\Sigma_{21} = \Sigma_{12} = 0$, what means $C_{21} = 0$. Finally, the matrix Q has to be considered diagonal.

Consequently, in a two scale volatility case, there is a first component of the volatility with a fast mean reversion and a second with a slow evolution. The introduction of a multiscale volatility allows to obtain a persistent skew for a medium maturity. In this case, we use a singular perturbations method and not a regular perturbation as in the case of a slow volatility time scale. Indeed, the parameters cannot be taken as zero with the perturbation in $\frac{1}{\varepsilon}$. The method remains the same but the calculations are much more complicated. The most important steps of the procedure as well as the important results are described but

some intermediate calculations are left to the reader.

The first Riccati equation can be studied under the form

$$\begin{aligned} A'(\tau) &= \frac{\theta(\theta-1)}{2} I_n + \frac{1}{\varepsilon} (-A(\tau)M_1 - M_1A(\tau) + 2A(\tau)Q_1^2A(\tau)) \\ &\quad + 2\theta \frac{1}{\sqrt{\varepsilon}} R^\top Q_1 A(\tau) + 2\theta \sqrt{\delta} R^\top Q_2 A(\tau) \\ &\quad + \delta (-A(\tau)M_2 - M_2A(\tau) + 2A(\tau)Q_2^2A(\tau)), \end{aligned}$$

with the following notations

$$\begin{aligned} M_1 &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad M = -\frac{1}{\varepsilon} M_1 - \delta M_2, \\ Q &= \frac{1}{\sqrt{\varepsilon}} Q_1 + \sqrt{\delta} Q_2, \quad Q_1 = \nu_1 M_1, \quad Q^2 = \nu_2 M_2. \end{aligned}$$

The development of the function A at order (ε, δ) is given by

$$A(\tau) = A^{0,0}(\tau) + \sqrt{\varepsilon} A^{1,0}(\tau) + \sqrt{\delta} A^{0,1}(\tau) + \varepsilon A^{2,0}(\tau) + \delta A^{0,2}(\tau) + \sqrt{\varepsilon} \sqrt{\delta} A^{1,1}(\tau) + o(\max(\varepsilon, \delta)).$$

The first steps of the procedure are detailed but the reader has to carry on some calculations. Indeed, contrary to the case of a long volatility time scale, the absence of a closed formula for the functions $A^{i,j}(\tau)$ requires an arduous identification step by step

Order $\frac{1}{\varepsilon}$:

$$0 = -A^{0,0}(\tau)M_1 - M_1A^{0,0}(\tau) + 2A^{0,0}(\tau)Q_1^2A^{0,0}(\tau).$$

Hence, $A^{0,0}(\tau)$ the form can be deduced

$$A^{0,0}(\tau) = \begin{pmatrix} 0 & 0 \\ 0 & a_{2,2}^{0,0}(\tau) \end{pmatrix}.$$

Order $\frac{1}{\sqrt{\varepsilon}}$:

$$\begin{aligned} 0 &= -A^{1,0}(\tau)M_1 - M_1A^{1,0}(\tau) + 2A^{1,0}(\tau)Q_1^2A^{0,0}(\tau) + 2A^{0,0}(\tau)Q_1^2A^{1,0}(\tau) \\ &\quad + 2\theta R^\top Q_1 A^{0,0}(\tau). \end{aligned}$$

Since $A^{0,0}(\tau)Q_1 = Q_1A^{0,0}(\tau) = 0$, the equation becomes $0 = A^{1,0}(\tau)M_1 + M_1A^{1,0}(\tau)$. Therefore, $A^{1,0}(\tau)$ is given by

$$A^{1,0}(\tau) = \begin{pmatrix} 0 & 0 \\ 0 & a_{2,2}^{1,0}(\tau) \end{pmatrix}.$$

Order 1 :

$$\begin{aligned} (A^{0,0})'(\tau) &= -A^{2,0}(\tau)M_1 - M_1A^{2,0}(\tau) + 2A^{2,0}(\tau)Q_1^2A^{0,0}(\tau) + 2A^{0,0}(\tau)Q_1^2A^{2,0}(\tau) \\ &\quad + 2A^{1,0}(\tau)Q_1^2A^{1,0}(\tau) + 2\theta R^\top Q_1 A^{1,0}(\tau) \\ &\quad + \frac{\theta(\theta-1)}{2} I_2. \end{aligned}$$

By noticing that $A^{0,0}(\tau)Q_1 = Q_1A^{0,0}(\tau) = A^{1,0}(\tau)Q_1 = Q_1A^{1,0}(\tau) = 0$, one can deduce

$$A^{0,0}(\tau) = \begin{pmatrix} 0 & 0 \\ 0 & \frac{\theta(\theta-1)}{2}\tau \end{pmatrix}, \quad A^{2,0}(\tau) = \begin{pmatrix} \frac{\theta(\theta-1)}{4} & 0 \\ 0 & a_{2,2}^{2,0}(\tau) \end{pmatrix}.$$

Step by step, all the coefficients $A^{i,j}(\tau)$ can be calculated. The constants obtained by integration are considered equal to 0 given that the global function A satisfies $A(0) = 0$. For information, the results of the procedure at order 2 are given so as to carry on the approximation procedure and to give a benchmark to the reader.

$$\begin{aligned} A^{0,0}(\tau) &= \begin{pmatrix} 0 & 0 \\ 0 & \frac{\theta(\theta-1)}{2}\tau \end{pmatrix}, \quad A^{1,0}(\tau) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad A^{0,1}(\tau) = \begin{pmatrix} 0 & 0 \\ 0 & \frac{\theta^2(\theta-1)}{2}R_{22}\nu_2\tau^2 \end{pmatrix}, \\ A^{2,0}(\tau) &= \begin{pmatrix} \frac{\theta(\theta-1)}{4} & 0 \\ 0 & 0 \end{pmatrix}, \quad A^{0,2}(\tau) = \begin{pmatrix} 0 & 0 \\ 0 & \frac{\theta^3(\theta-1)}{3}R_{22}^2\nu_2^2\tau^3 - \frac{\theta(\theta-1)}{2}\tau^2 + \frac{\theta^2(\theta-1)^2}{6}\nu_2^2\tau^3 \end{pmatrix}, \\ A^{1,1}(\tau) &= \begin{pmatrix} 0 & 0 \\ 0 & \theta^3(\theta-1)R_{12}R_{21}\nu_1\nu_2\tau^2 \end{pmatrix}, \quad A^{2,1}(\tau) = \begin{pmatrix} 0 & \theta^2(\theta-1)R_{21}\nu_2\tau \\ 0 & * \end{pmatrix}, \\ A^{3,0}(\tau) &= \begin{pmatrix} \frac{\theta^2(\theta-1)}{4}R_{11}\nu_1 & 0 \\ \frac{\theta^2(\theta-1)}{2}R_{12}\nu_1 & * \end{pmatrix}, \quad A^{3,1}(\tau) = \begin{pmatrix} 0 & 2\theta^3R_{11}\nu_1\nu_2R_{21}(\theta-1)\tau \\ 0 & * \end{pmatrix}, \\ A^{4,0}(\tau) &= \begin{pmatrix} \frac{\theta^2(\theta-1)^2}{16}\nu_1^2 + \frac{\theta^3(\theta-1)}{4}R_{11}^2\nu_1^2 & 0 \\ \frac{\theta^3(\theta-1)}{2}R_{11}\nu_1^2R_{12} & * \end{pmatrix}. \end{aligned}$$

Concerning the function C , the third Riccati equation allows to obtain the following development

$$C(\tau) = C^{0,0}(\tau) + \sqrt{\varepsilon}C^{1,0}(\tau) + \sqrt{\delta}C^{0,1}(\tau) + \varepsilon C^{2,0}(\tau) + \delta C^{0,2}(\tau) + \sqrt{\varepsilon}\sqrt{\delta}C^{1,1}(\tau) + o(\max(\varepsilon, \delta)),$$

where

$$\begin{aligned} C^{0,0}(\tau) &= r\tau\theta + \beta \int_0^\tau \text{Tr}(Q_1^2 A^{20}(s))ds = r\tau\theta + \frac{\theta(\theta-1)}{4}\beta\nu_1^2\tau. \\ C^{1,0}(\tau) &= \beta \int_0^\tau \text{Tr}(Q_1^2 A^{30}(s))ds = \frac{\theta^2(\theta-1)}{4}\beta\nu_1^3R_{11}\tau. \\ C^{0,1}(\tau) &= \beta \int_0^\tau \text{Tr}(Q_1^2 A^{21}(s))ds = 0. \\ C^{2,0}(\tau) &= \beta \int_0^\tau \text{Tr}(Q_1^2 A^{40}(s))ds = \frac{\theta^2(\theta-1)^2}{16}\beta\nu_1^4\tau + \frac{\theta^3(\theta-1)}{4}\beta\nu_1^4R_{11}^2\tau. \\ C^{0,2}(\tau) &= \beta \int_0^\tau \text{Tr}(Q_2^2 A^{00}(s))ds = \frac{\theta(\theta-1)}{4}\beta\nu_2^2\tau^2. \\ C^{1,1}(\tau) &= \beta \int_0^\tau \text{Tr}(Q_1^2 A^{31}(s))ds = 0. \end{aligned}$$

Call price and smile approximation

As before, the Fourier transform expression can be calculated

$$\Phi_t(T, \theta) = \exp \left[\frac{\theta(\theta-1)}{2} (T-t) \left(\frac{\beta \nu_1^2}{2} + V_t^{22} \right) + \theta y + r(T-t)\theta \right] * \left[\begin{array}{l} 1 + \sqrt{\varepsilon} [Tr(A^{1,0}V_t) + C^{1,0}] + \sqrt{\delta} Tr(A^{0,1}V_t) \\ + \varepsilon [Tr(A^{2,0}V_t) + C^{2,0} + \frac{1}{2} ((Tr(A^{1,0}V_t) + C^{1,0})^2)] \\ + \delta [Tr(A^{0,2}V_t) + C^{0,2} + \frac{1}{2} (Tr(A^{0,1}V_t))^2] \\ + \sqrt{\varepsilon}\sqrt{\delta} [Tr(A^{1,1}V_t) + (Tr(A^{1,0}V_t) + C^{1,0}) Tr(A^{0,1}V_t)] \\ + o(\max(\varepsilon, \delta)) \end{array} \right],$$

with

$$\begin{aligned} Tr(A^{1,0}V_t) &= 0. \\ Tr(A^{0,1}V_t) &= \frac{\theta^2(\theta-1)}{2} R_{22}\nu_2 V_t^{22} (T-t)^2. \\ Tr(A^{1,1}V_t) &= \theta^3(\theta-1) R_{21} R_{12} \nu_2 \nu_1 V_t^{22} (T-t)^2. \\ Tr(A^{2,0}V_t) &= \frac{\theta(\theta-1)}{4} V_t^{11}. \\ Tr(A^{0,2}V_t) &= \frac{\theta^3(\theta-1)}{3} R_{22}^2 \nu_2^2 V_t^{22} (T-t)^3 - \frac{\theta(\theta-1)}{2} V_t^{22} (T-t)^2 + \frac{\theta^2(\theta-1)^2}{6} \nu_2^2 V_t^{22} (T-t)^3. \end{aligned}$$

In regard to the smile for a slow volatility time scale, the same procedure allows to find the expected approximation. Moreover, this method is easier to implement than the standard singular perturbations method [3] and an expansion at order 2 is given. The premium can be developed as follows

$$P = P^{0,0} + \sqrt{\varepsilon} P^{1,0} + \sqrt{\delta} P^{0,1} + \varepsilon P^{2,0} + \delta P^{0,2} + \sqrt{\varepsilon}\sqrt{\delta} P^{1,1} + o(\max(\varepsilon, \delta)),$$

with

$$\begin{aligned} P^{0,0} &= C_{BS(\sigma)}, \quad \sigma = \sqrt{\frac{\beta \nu_1^2}{2} + V_t^{22}}. \\ P^{1,0} &= \frac{\beta R_{11} \nu_1^3 (T-t)}{4} \Lambda_y^{BS,3}, \quad P^{0,1} = \frac{R_{22} \nu_2 (T-t)^2 V_t^{22}}{2} \Lambda_y^{BS,3}. \\ P^{2,0} &= \frac{1}{4} V_t^{11} \Lambda_y^{BS,2} + \beta \nu_1^4 (T-t) \frac{1}{16} (\Lambda_y^{BS,4} - \Lambda_y^{BS,3}) + \frac{1}{4} \beta \nu_1^4 R_{11}^2 (T-t) \Lambda_y^{BS,4} \\ &\quad + \frac{1}{2} \frac{\beta^2 R_{11}^2 \nu_1^6 (T-t)^2}{16} (\Lambda_y^{BS,6} - \Lambda_y^{BS,5}). \\ P^{0,2} &= \frac{1}{4} \beta \nu_2^2 (T-t)^2 \Lambda_y^{BS,2} - \frac{1}{2} V_t^{22} (T-t)^2 \Lambda_y^{BS,2} + \frac{1}{3} R_{22}^2 \nu_2^2 V_t^{22} (T-t)^3 \Lambda_y^{BS,4} \\ &\quad + \frac{1}{6} \nu_2^2 V_t^{22} (T-t)^3 (\Lambda_y^{BS,4} - \Lambda_y^{BS,3}) + \frac{1}{2} \frac{R_{22}^2 \nu_2^2 (T-t)^4 (V_t^{22})^2}{4} (\Lambda_y^{BS,6} - \Lambda_y^{BS,5}). \\ P^{1,1} &= R_{21} R_{12} \nu_2 \nu_1 V_t^{22} (T-t)^2 \Lambda_y^{BS,4} + \frac{\beta R_{11} \nu_1^3 (T-t)^3}{4} \frac{R_{22} \nu_2 V_t^{22}}{2} (\Lambda_y^{BS,6} - \Lambda_y^{BS,5}). \end{aligned}$$

As it was done before for the long volatility time scale, by injecting the Taylor expansion of the implied volatility for Call options in the Black-Scholes model, and by comparing both approximations of the price, an approximation for the implied volatility at order (ε, δ) is deduced

$$\begin{aligned}
 \Sigma^{0,0} &= \sqrt{\frac{\beta\nu_1^2}{2} + V_t^{22}}. \\
 \Sigma^{1,0} &= \frac{P^{1,0}}{Vega^{BS}} = -\frac{1}{\sqrt{T-t}(\frac{\beta\nu_1^2}{2} + V_t^{22})} \frac{\nu_1^3}{4} \beta R_{11} d_0. \\
 \Sigma^{0,1} &= \frac{P^{0,1}}{Vega^{BS}} = -\frac{1}{\sqrt{T-t}(\frac{\beta\nu_1^2}{2} + V_t^{22})} \frac{\nu_2}{2} R_{22} V_t^{22} d_0 \\
 \Sigma^{2,0} &= \frac{P^{2,0} - \frac{1}{2} (\Sigma^{1,0})^2 Vomma^{BS}}{Vega^{BS}}. \\
 \Sigma^{0,2} &= \frac{P^{0,2} - \frac{1}{2} (\Sigma^{0,1})^2 Vomma^{BS}}{Vega^{BS}}. \\
 \Sigma^{1,1} &= \frac{P^{1,1} - \Sigma^{1,0} \Sigma^{0,1} Vomma^{BS}}{Vega^{BS}}.
 \end{aligned}$$

The approximation is available at order (ε, δ) but for a better readability, we will present only the expression of the smile at $(\sqrt{\varepsilon}, \sqrt{\delta})$

$$\begin{aligned}
 \hat{\Sigma}_t(T, K) &= \sqrt{\frac{\beta\nu_1^2}{2} + V_t^{22}} + \frac{\sqrt{\varepsilon} P^{1,0}(t, Y_t, V_t^{22}) + \sqrt{\delta} P^{0,1}(t, Y_t, V_t^{22})}{\frac{\partial C_t^{BS}}{\partial \sigma}(T, K, \sqrt{\frac{\beta\nu_1^2}{2} + V_t^{22}})} \\
 &= \sqrt{\frac{\beta\nu_1^2}{2} + V_t^{22}} \\
 &\quad + \frac{1}{(\frac{\beta\nu_1^2}{2} + V_t^{22})^{\frac{3}{2}}} \left[\frac{q_1^3}{4m_1^2} \beta R_{11} + \frac{q_2}{2} R_{22} (T-t) V_t^{22} \right] \left[\frac{\log(K e^{-r(T-t)}) - Y_t}{(T-t)} + \frac{\frac{\beta\nu_1^2}{2} + V_t^{22}}{2} \right].
 \end{aligned}$$

For a two-scale volatility, the implied volatility $\Sigma^{0,0} = \sqrt{\frac{\beta\nu_1^2}{2} + V_t^{22}}$ at order 0 is a combination of the long term and the short term through the stationary value of V_s^{11} and the initial value of V_s^{22} . Moreover, the skew around the forward money ($K = F_t^T = e^{Y_t + r(T-t)}$) can also be inferred at order $(\sqrt{\varepsilon}, \sqrt{\delta})$. The analysis is focused on the skew in order to give characteristics of the model related to the stochastic correlation. But it is obvious that the same reasoning can be done for convexity and slope.

$$\begin{aligned}
 \frac{\partial \Sigma_t}{\partial K}(T, F_t^T) &\sim \frac{1}{F_t^T} \frac{1}{(\frac{\beta\nu_1^2}{2} + V_t^{22})^{\frac{3}{2}}} \left[\frac{q_1^3}{4m_1^2(T-t)} \beta R_{11} + \frac{q_2}{2} R_{22} V_t^{22} \right] \\
 &\sim \frac{1}{F_t^T} \frac{1}{(\frac{\beta\nu_1^2}{2} + V_t^{22})^{\frac{3}{2}}} \left[\nu_1^3 \frac{\sqrt{\varepsilon}}{T-t} \beta R_{11} + \frac{\nu_2}{2} \sqrt{\delta} R_{22} V_t^{22} \right]. \quad (2.24)
 \end{aligned}$$

This formula underlines that the *skew* splits up into two components : the first one, proportional to R_{11} , coming from the fast mean-reversion volatility component, and the second one, proportional to R_{22} , is persistent and arises from the slow variation component of the volatility. Besides, this expression is compatible with the combination of Heston formulas for short and long maturities [2].

2.3.6 Numerical Applications

This part is devoted to numerical applications. Indeed, this section lays the emphasis on the error of the approximation as well as the influence of some parameters like the matrix R and the matrix Q on the smile. It is noticeable that for a range of parameters (a set of parameters in a classic market), this approximation can be used for calibration. But when the parameters overstep some bounds, the approximation does not furnish a calibration tool but carries on providing an idea of the level and the shape of the smile as well as the influence of parameters.

Smile for a slow volatility time scale

Let us handle now the case of a slow time volatility scale compared with the maturity option i.e $(T - t) \ll \tau_1, \tau_2$. At $t = 0$, consider the followings values for the parameters of the model allowing standard dynamics for the volatility in short maturities

$$S_0 = 100, \quad r = 0, \quad \beta = 3.6, \quad V_0 = \begin{pmatrix} 0.02 & 0.01 \\ 0.01 & 0.02 \end{pmatrix}, \quad M = \begin{pmatrix} -0.05 & 0 \\ 0 & -0.05 \end{pmatrix}, \quad Q = \begin{pmatrix} 0.02 & 0.01 \\ 0.01 & 0.02 \end{pmatrix}.$$

The matrix of mean-reversion M gives the speed at which the volatility process returns to its mean. The matrix Q corresponds to the volatility of volatility which is the most significant parameter of the stochastic volatility model concerning the evolution of the volatility and the dispersion of the volatility around its expected value. First, the matrix Q was chosen so that the volatility does not reach quickly the asymptotic implied volatility ATM $\Sigma_\infty \sim \sqrt{\frac{\beta}{2}(\nu_1^2 + \nu_2^2)} \sim 17,8\%$ whereas $\Sigma_0 = \sqrt{V_0^{11} + V_0^{22}} = 20\%$. This case is common in a classic market.

Numerical applications underline the effect of the correlation matrix R on a smile for a maturity $T = 6$ months ($6m$). Indeed, when R is diagonal, parameters can be found so that the Wishart volatility model corresponds to a multifactor Heston model described for example in [7]. Then, adding non-diagonal components for R allows more flexibility concerning the stochastic correlation and the skew because they will also depend on the new factor V_t^{12} . The study is restricted to the case when M and Q are diagonals because it is the framework of this approximation, nevertheless a general matrix R can be considered in order to glimpse the flexibility of the Wishart volatility model. The impact of the non-diagonal components for M and Q will be studied in future researches.

Characteristic times of the process V are denoted by $\tau_1 = \frac{1}{2m_1} = 10$ years ($10y$) and $\tau_2 = \frac{1}{2m_2} = 10$ years. Then, this framework deals with a "short maturity" case. At $t = 0$, we have calculated the smile by FFT (Figure 2.1) and the approached smiles at order 1 and order 2 (Figure 2.2 and Figure 2.3). The implied volatility level is represented on the Y-axis and the strike is represented on the X-axis.

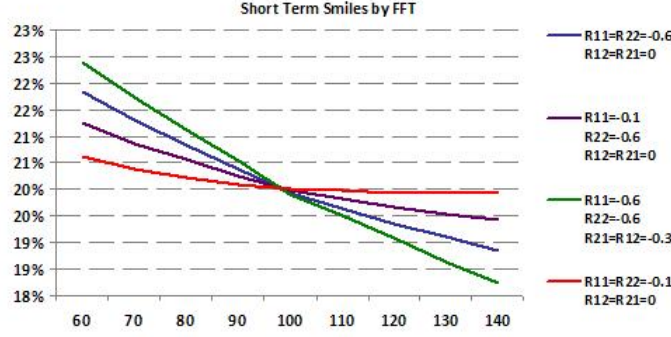


FIGURE 2.1 – Short term smile by FFT ($T = 6m \ll \tau_1 = \tau_2 = 10y$)

As expected, all components of the correlation matrix R contribute for the generation of the skew. Indeed, the more the components R_{11} and R_{22} are negative, the more the skew is negative. That is not true for R_{12} and R_{21} since they are multiplied by V_0^{12} and V_0^{21} whose are not necessarily positive.

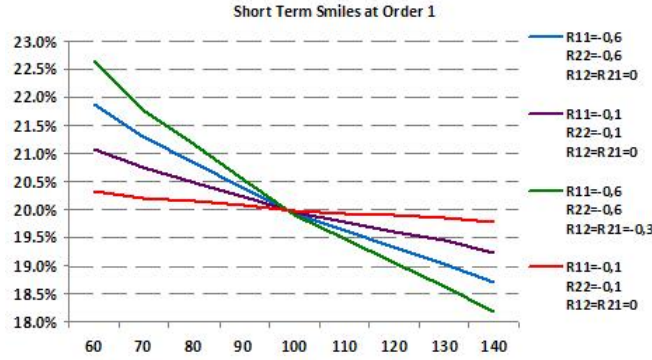


FIGURE 2.2 – Short term smile at order $(\sqrt{\varepsilon}, \sqrt{\delta})$ ($T = 6m \ll \tau_1 = \tau_2 = 10y$)

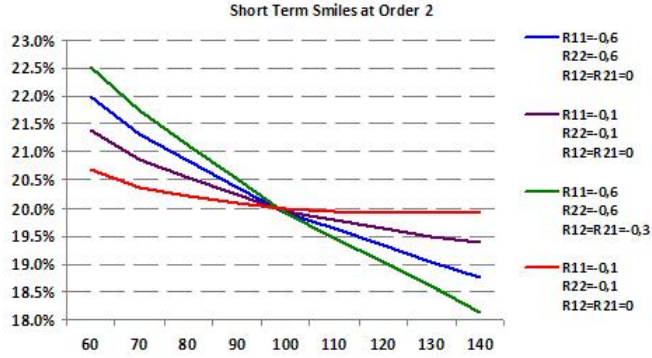


FIGURE 2.3 – Short term smile at order (ε, δ) ($T = 6m \ll \tau_1 = \tau_2 = 10y$)

In order to quantify the accuracy of this approximation, figures are presented (Figure 2.4 and Figure 2.5) describing the difference between the approached smiles and the smile obtained by FFT.

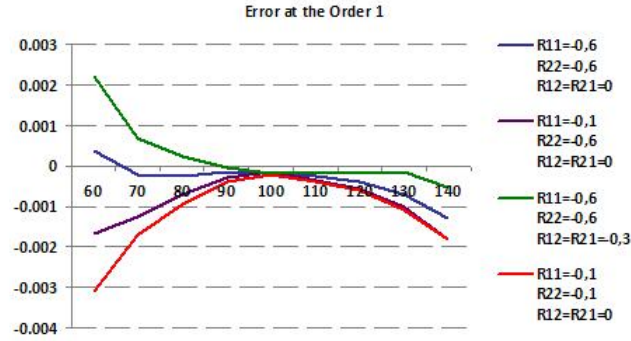


FIGURE 2.4 – Error for short term smile at order $(\sqrt{\varepsilon}, \sqrt{\delta})$ ($T = 6m \ll \tau_1 = \tau_2 = 10y$)

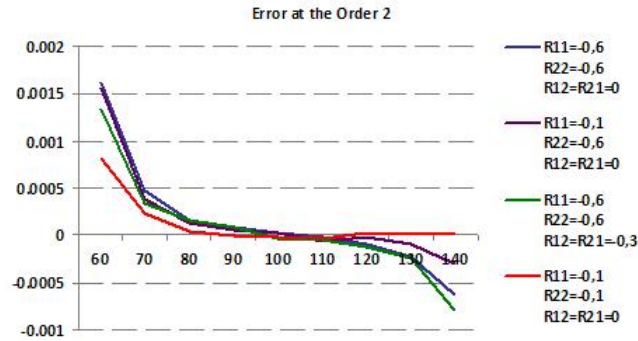


FIGURE 2.5 – Error for short term smile at order (ε, δ) ($T = 6m \ll \tau_1 = \tau_2 = 10y$)

The approximation gives very good results and the approximation at order 2 is accurate enough to be used as a calibration tool : one can see that between the strike 80 and 120, the error is below 5×10^{-5} . This error is totally acceptable by practitioners in a calibration prospect.

An unrealistic case is presented where the volatility of volatility matrix Q is outstanding so that the volatility explodes when maturity tends to infinity. Although M is small what implied a slow mean reversion, the level of the implied volatility can be increased by the volatility of volatility Q . Indeed, as seen in Comment 2.3, the level of the asymptotic volatility depends on QQ^T . Only the matrix Q is changed to see its influence on the smile and the behaviour of the approximation in a "tense" framework. With these market conditions, the approximation is not good enough in a calibration prospect, but keeps on giving a good shape and level of the smile and consequently a good idea of the impact of the parameters.

$$S_0 = 100, \quad r = 0, \quad \beta = 3.6, \quad V_0 = \begin{pmatrix} 0.02 & 0.01 \\ 0.01 & 0.02 \end{pmatrix}, \quad M = \begin{pmatrix} -0.05 & 0 \\ 0 & -0.05 \end{pmatrix}, \quad Q = \begin{pmatrix} 0.1 & 0.05 \\ 0.05 & 0.1 \end{pmatrix}.$$

In this case, a great asymptotic implied volatility ATM ($\Sigma_\infty \sim \sqrt{\frac{\beta}{2}(\nu_1^2 + \nu_2^2)} \sim 89\%$) is considered and the smile is wealthier as it can be seen in Figure 2.6. First, the level at the money (ATM) of the implied volatility at $T = 6$ months is about 25% and is then very different from the initial implied volatility which is $\Sigma_0 = \sqrt{V_0^{11} + V_0^{22}} = 20\%$. Moreover, one can obtain a real "smirk" with these parameters, in the fourth case for instance (Figure 2.6). The Black-Scholes volatility used in the FFT pricing with Formula (2.19) is $\bar{\sigma} = 24.22\%$ which is the realized mean volatility giving an account of the real level of the volatility ATM.

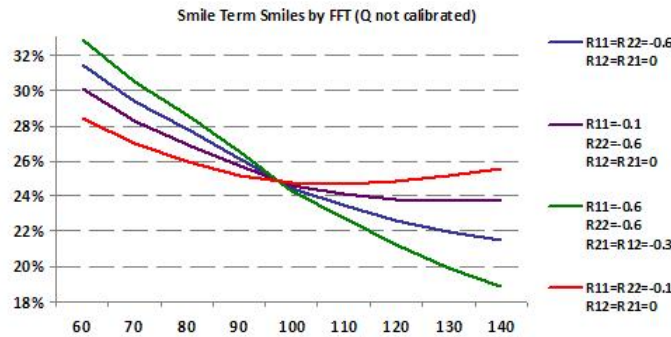


FIGURE 2.6 – Short term smile by FFT ($T = 6m \ll \tau_1 = \tau_2 = 10y$)

Although the calibration with this approximation is impossible for those parameters, the approximation appears to be stable enough in relation to shock on parameters and

reproduces a good shape for the smile like the level ATM, the skew and a good behaviour of the convexity.

The approximation allows to have an idea of the smile and the influence of some parameters (very useful for a trader) but cannot furnish a calibration tool. For the 4 sets of correlation coefficients considered, comparing the smile by FFT and the smiles approached by the method at order 1 and 2 allows to estimate the convergence of this method.

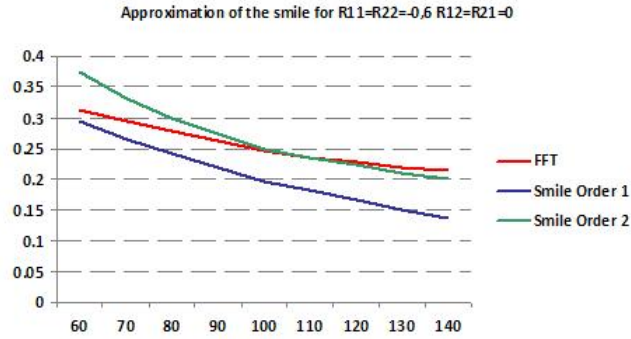


FIGURE 2.7 – Convergence in the case : $R_{11} = R_{22} = -0,6$ $R_{12} = R_{21} = 0$

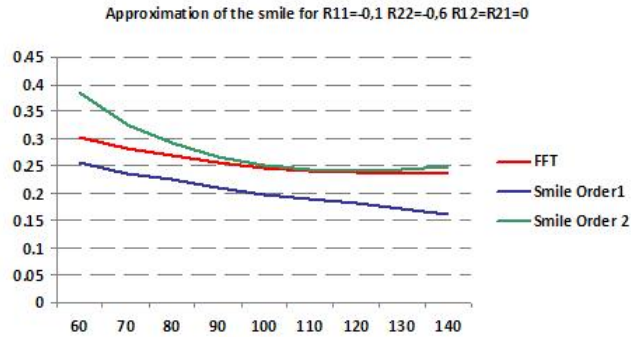
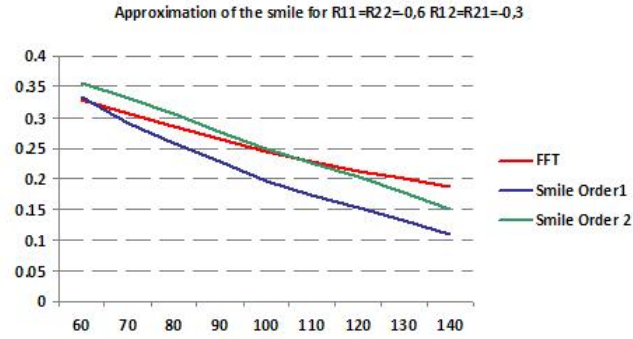
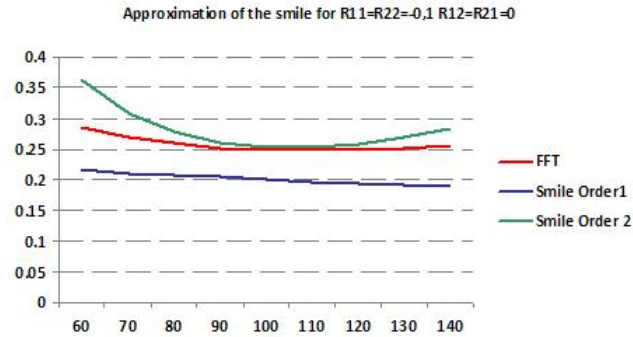


FIGURE 2.8 – Convergence in the case : $R_{11} = R_{22} = -0,1$ $R_{12} = R_{21} = 0$


 FIGURE 2.9 – Convergence in the case : $R_{11} = R_{22} = -0.6$ $R_{12} = R_{21} = -0.3$

 FIGURE 2.10 – Convergence in the case : $R_{11} = -0.1$ $R_{22} = -0.6$ $R_{12} = R_{21} = 0$

The perturbation method becomes more accurate at order 2 : the approximation reaches the good level ATM and the shape of the curves is reproduced more precisely. Indeed, the smirk in the fourth case is faithfully reproduced and the smile level ATM is not far from the one obtained by FFT (24% see Figure 2.6).

For information, a complete surface of a short term smile is presented functions of maturity in years and strike (Figure 2.11). The slope appears very high and the smile does not flatten itself when the maturity increases in opposition to the case of a local volatility model.

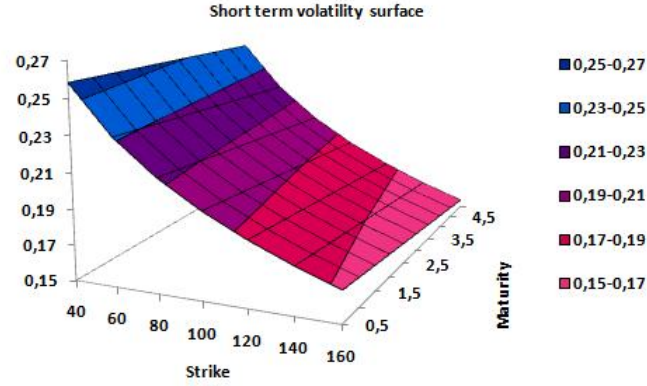


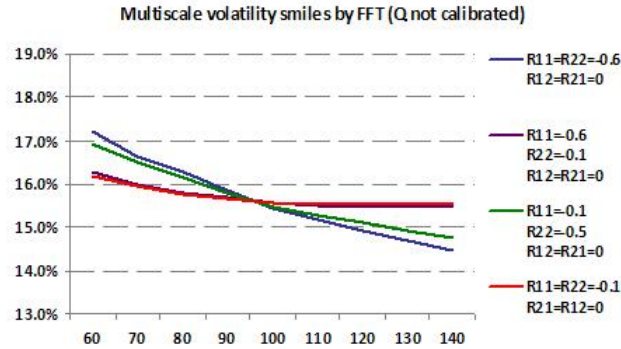
FIGURE 2.11 – Implied volatility surface for short maturities

Smile with multiscale volatility

At present, those new values for the model parameters are taken as input

$$S_0 = 100, \quad r = 0, \quad \beta = 3.6, \quad V_0 = \begin{pmatrix} 0.02 & 0.01 \\ 0.01 & 0.02 \end{pmatrix}, \quad M = \begin{pmatrix} -2 & 0 \\ 0 & -0.02 \end{pmatrix}, \quad Q = \begin{pmatrix} 0.05 & 0 \\ 0 & 0.02 \end{pmatrix}.$$

Consider a maturity $T = 3y$ what allows the case of a medium maturity with $\tau_1 = 3m \ll T = 3y \ll \tau_2 = 25y$. The matrix Q is chosen so that the implied volatility has gone from $\Sigma_0 = 0\%$ until $\Sigma_{3Y} \sim \sqrt{\beta \frac{\nu_1^2}{2} + V_0^{22}} \sim 15.6\%$ in 3 years (Figure 2.12) : it is an significative change of implied volatility and the quality of the method has to be tested seeing how the approximation converges toward the real smile with these exceptional market conditions. At order 1, the perturbation method furnishes a good shape for the skew but the level of the smile at the money is far from the good one (14.9% see Figure 2.13). However, at order 2, the approximated smile is accurate and the level is really close to the good one (15.7% see Figure 2.14).


 FIGURE 2.12 – Multiscale volatility smiles by FFT ($\tau_1 = 3m \ll T = 3y \ll \tau_2 = 25y$)

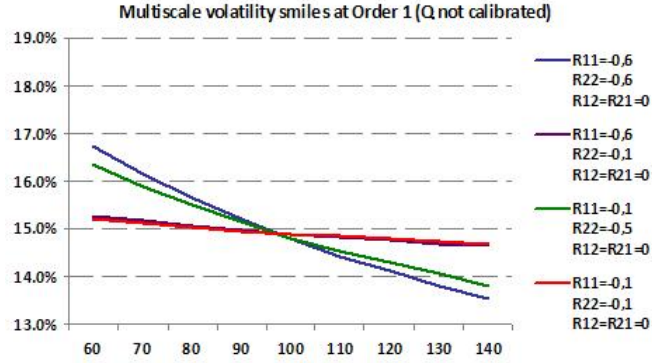


FIGURE 2.13 – Multiscale volatility smiles at order $(\sqrt{\varepsilon}, \sqrt{\delta})$ ($\tau_1 = 3m \ll T = 3y \ll \tau_2 = 25y$)

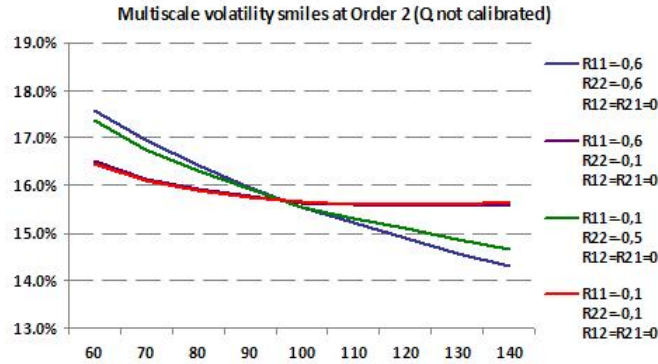


FIGURE 2.14 – Multiscale volatility smiles at order (ε, δ) ($\tau_1 = 3m \ll T = 3y \ll \tau_2 = 25y$)

It is important to point out that the approximation may be not accurate enough for calibration because the set of parameters was taken such as the volatility significantly changes. Moreover, as in the short maturity case, the procedure gives a good idea of the smile as well as the influence of some parameters (very useful for traders). The smile by FFT is approached by the method at order 1 and at order 2 for the same sets of correlation coefficients (Figure 2.15 and Figure 2.16).



FIGURE 2.15 – Error for multiscale volatility smile at order $(\sqrt{\varepsilon}, \sqrt{\delta})$ ($\tau_1 = 3m \ll T = 3y \ll \tau_2 = 25y$)



FIGURE 2.16 – Error for multiscale volatility smile at order (ε, δ) ($\tau_1 = 3m \ll T = 3y \ll \tau_2 = 25y$)

As expected, the skew comes from the component R_{22} and the slope of the smile curve is almost insensible to a change of R_{11} given that $\tau_1 \ll T$. This is what it was expected after an analysis of the smile through the approximation formula (2.24). We can also describe the implied volatility surface in the case of a medium maturity option (see Figure 2.17).

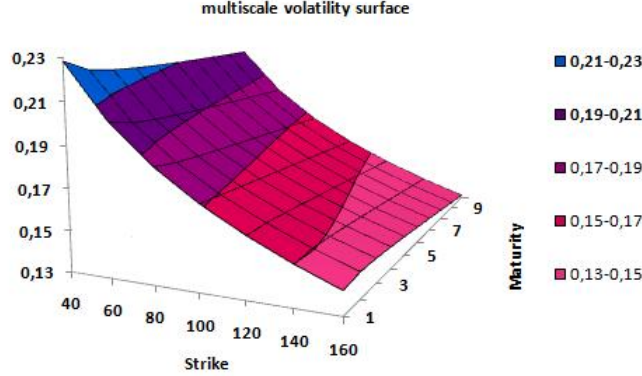


FIGURE 2.17 – Implied volatility surface for medium maturities

2.4 Study of the influence of the parameters

In order to evaluate the performances of the model, we describe the influence of the parameters on the smile. Considering call options, we analyze the possible shapes of smiles generated by the model and how smiles can be managed by the parameters. The non diagonal components of the matrices are specific to the Wishart model and we will the great importance of these parameters.

For the test of the influence of the parameters, we have chosen to use the approximation at order 2 which is a closed formula so that there is no numerical error coming from an evaluation by Monte Carlo or other numerical methods. Moreover, these tests allow to understand the behavior of the approximation with report to the parameters. It is complicated to quantify the impact of the terms in the approximation at order 2. Indeed, for each term of the approximation of the price, there are different kind of components : there is a component proportional to successive derivatives of the logarithm of the underlying in the Black-Scholes model, another component proportional to parameters and a third component proportional to a power of the maturity. For the testing of the influence of the parameters, we will focus only on the component proportional to the parameters. In the price P^2 given by the formula (2.21), there is a term proportional to MV_0 , a term proportional to $Q^\top Q$, a term proportional to $Q^\top QV_0$, a term proportional to $R^\top QV_0$ and a term proportional to $(R^\top Q)^2V_0$. Therefore, the parameters Q and V_0 seem essential and we will study their influence on the smile. This influence is weighted by the correlation through the terms $R^\top QV_0$ and $(R^\top Q)^2V_0$.

In this section, we consider a Call option at maturity $T = 6m$ and we assume in the following that $S_0 = 100$ and $r = 0$. Moreover, the notation * underlines the parameters

studied concerning their influence on the smile.

2.4.1 Influence of the parameter β

The parameter β represents the degree of the Wishart process considering the link between the Wishart process and Ornstein-Uhlenbeck processes. Consider the following set of parameters :

$$\beta = *, V_0 = \begin{pmatrix} 0.02 & 0.01 \\ 0.01 & 0.02 \end{pmatrix}, M = \begin{pmatrix} -0.05 & 0 \\ 0 & -0.03 \end{pmatrix}, Q = \begin{pmatrix} 0.02 & 0.01 \\ 0.01 & 0.02 \end{pmatrix}, R = \begin{pmatrix} 0.6 & 0 \\ 0 & 0.6 \end{pmatrix}.$$

The influence of the parameter β can be seen in the Figure 2.18. This parameter has a direct influence on the level of the smile. However, its influence on the slope and on the convexity is very low. Therefore, this parameter will be essential in order to fit the smile in the calibration procedure.

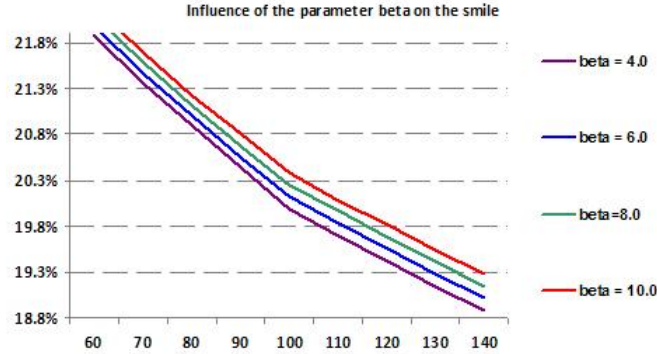


FIGURE 2.18 – Influence of the parameter β on the smile

2.4.2 Influence of the correlation R

The correlation matrix does not influence the level of the smile at the money but can manage the shape of the smile, particularly the sign and the value of the skew at the money. In a multidimensional Heston model, the stochastic correlation is a linear combination of the volatility factors using principally the correlation components. Then, the dynamics of the skew are managed by the correlation parameters. In the Wishart model, the constraint is not as strong as for the Heston model because of an additional degree of freedom for the stochastic correlation resulting from the non-diagonal component of the Wishart process defined at (2.8).

Moreover, the norm of the correlation has an impact on the smile. The weaker the norm of the correlation, the more convex is the smile. Indeed, a low correlation leads to a low skew what generally increases the convexity of the smile. When the matrix R is low, the convexity increases due to the influence of the matrix Q . On the contrary, when the correlation is high, the skew is important and the smile looks like a straight line. This can be interpreted as a tension effect that reduces the convexity when the slope is important.

2.4.3 Influence of the volatility of volatility Q

The matrix Q represents the volatility of the volatility and has a major impact on the smile structure. We analyze the influence of the diagonal components and the influence of the non diagonal components.

Influence of Q_{11} and Q_{22}

We have seen that the diagonal parameters of the matrix Q have an influence on the smile level. For very short maturities, the level of the smile is close to the value $Tr(V_0)$ that represents the initial stochastic volatility. In practice, for a slow volatility time scale, the level of the smile is different from this value : this difference increases when the diagonal components of the matrix Q increase.

Let consider the following set of parameters :

$$\beta = 4, V_0 = \begin{pmatrix} 0.02 & 0.01 \\ 0.01 & 0.02 \end{pmatrix}, M = \begin{pmatrix} -0.05 & 0 \\ 0 & -0.03 \end{pmatrix}, Q = 0.1 \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}.$$

The diagonal components of the matrix Q have an impact on the convexity of the smile. However, the convexity remains low even if the values of the components increase at very high levels (see Figure 2.19).

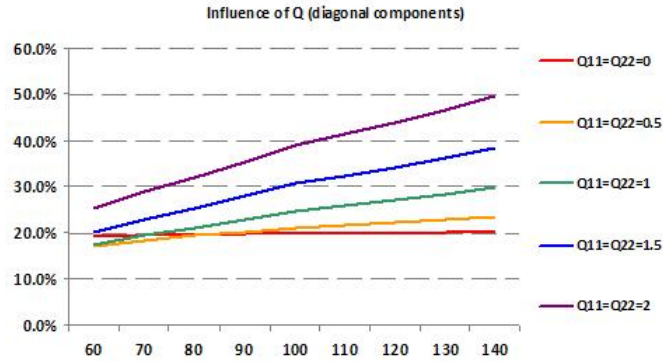


FIGURE 2.19 – Influence of the diagonal components of Q on the smile in the case $R = 0.6I_n$

Smile with convexity can be obtained when the correlation parameters are low enough. But the shape of the smile seem to be restricted and the convexity seems to be limited. We analyze if the non diagonal components of the matrix Q can improve the behavior of the model allowing various shapes of smile.

Influence of Q_{12} and Q_{21}

We apply the same idea that consists in fixing a set of parameters and in analyzing how the smile varies with a change of the parameters. Let consider the initial set of parameters when the non diagonal components of Q are null.

$$\beta = 4, V_0 = \begin{pmatrix} 0.02 & 0.01 \\ 0.01 & 0.02 \end{pmatrix}, M = \begin{pmatrix} -0.05 & 0 \\ 0 & -0.03 \end{pmatrix}, Q = \begin{pmatrix} 0.1 & * \\ * & 0.1 \end{pmatrix}.$$

We consider four sets of correlation such as the model can generate various shapes of smiles with different monotonies and convexities as can be seen in the figures from Figure 2.21 to Figure 2.22.

$$\begin{aligned} R_{11} = R_{22} = 0.6 \quad R_{12} = R_{21} = 0 \\ R_{11} = R_{22} = -0.6 \quad R_{12} = R_{21} = 0 \\ R_{11} = R_{22} = -0.6 \quad R_{12} = R_{21} = -0.6 \\ R_{11} = R_{22} = 0.6 \quad R_{12} = R_{21} = 0.6 \end{aligned}$$

Then, we change the non diagonal components of the matrix Q keeping the others parameters unchanged. We observe the evolution of the smiles depending also on the correlation. At order 1, the influence of the matrix Q is totally weighted by the correlation and at order 2, there are different terms depending on Q , R and M as seen in (2.21).

When the matrix Q is high, the terms in the price at order 2 that depend only on Q are preponderant comparing with other terms for all the strikes between 70 and 130. At order 1, there is no influence of Q without a weighting of R but at order 2 there is an influence of the parameter Q . The term proportional to $Tr(MV_0)$ is not really important if we consider short maturities.

First tests : case of a symmetric matrix Q

We consider the influence of the non diagonal components of Q with a correlation fixed at $R = 0.6I_n$. We remark that their influence is important on the level and the convexity. Moreover, the influence on convexity is much more important with these components (see Figure 2.20) than with diagonal components (see Figure 2.19).

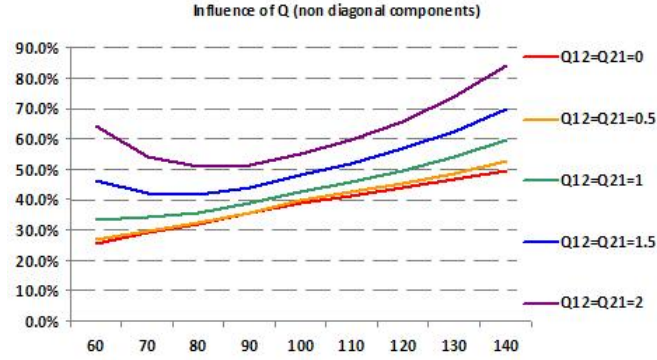


FIGURE 2.20 – Influence of the non diagonal components of Q on the smile in the case $R = 0.6I_n$

We consider the evolution of the smiles for different sets of correlation when the non diagonal coefficients change.

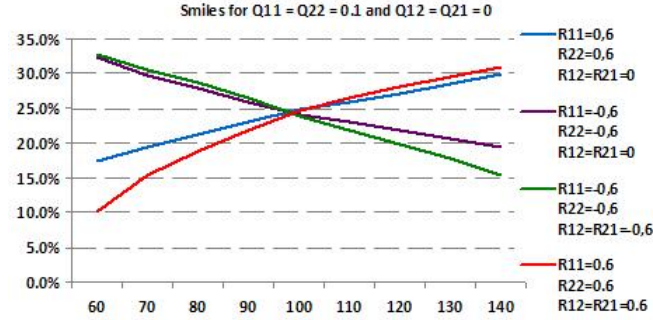


FIGURE 2.21 – Smiles for $Q_{11} = Q_{22} = 0.1$ and $Q_{12} = Q_{21} = 0$

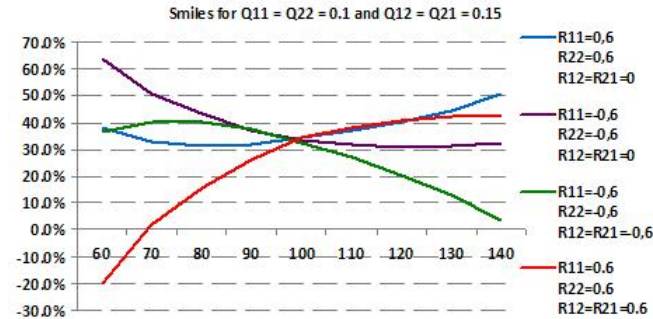


FIGURE 2.22 – Smiles for $Q_{11} = Q_{22} = 0.1$ and $Q_{12} = Q_{21} = 0.15$

One can notice that the non diagonal coefficients of Q have an important influence on the level and the shape of the smile. Indeed, the diagonal components does not provide wealthy shapes whatever the set of correlations. However, the non diagonal components generates convexity. For example, we can compare the influence of the coefficients Q by considering the smiles in two cases : the first one with $Q_{11} = Q_{22} = 0.1$ and $Q_{12} = Q_{21} = 0$ and the second one with $Q_{11} = Q_{22} = 0$ and $Q_{12} = Q_{21} = 0.1$.

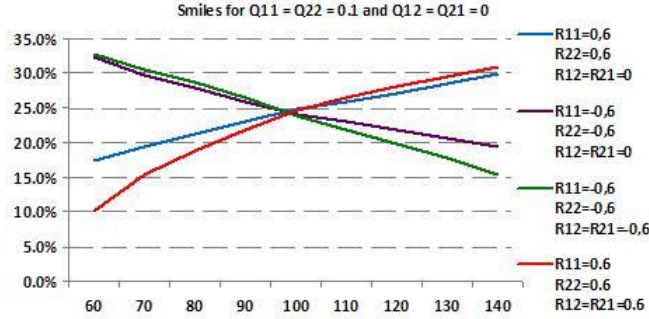


FIGURE 2.23 – Smiles for $Q_{11} = Q_{22} = 0.1$ and $Q_{12} = Q_{21} = 0$

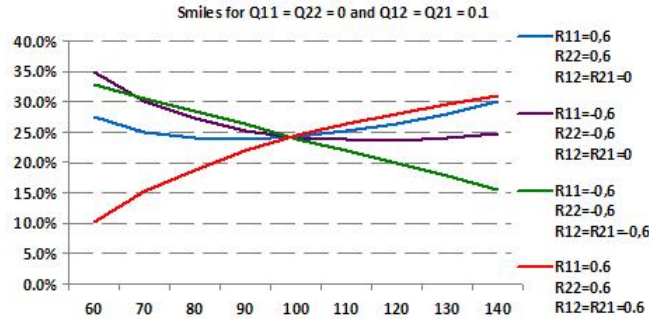


FIGURE 2.24 – Smiles for $Q_{11} = Q_{22} = 0$ and $Q_{12} = Q_{21} = 0.1$

The Wishart model is flexible and can generate various shapes of smile with convexity and change of monotony. It could be interesting to find a market such as forward market and FOREX where the Wishart model succeeds in capturing phenomena whereas other models fail [11].

General tests : case of a full matrix Q

The study we have done gives an intuition but does not gives a clear interpretation of the influence of the parameter of Q . We aim at finding a factor depending on Q that controls

only the convexity and not the level of the smile. The matrix $Q^\top Q$ has interesting properties. We recall the dynamic of the Wishart model :

$$\begin{aligned}\frac{dS_t}{S_t} &= rdt + \sqrt{Tr(V_t)}dB_t \\ dTr(V_t) &= (\beta Tr(Q^\top Q) + 2Tr(MV_t))dt + 2\sqrt{Tr(Q^\top QV_t)}(\rho_t dB_t + \sqrt{1 - \rho_t^2}d\tilde{B}_t) \\ \rho_t &= \frac{Tr(R^\top QV_t)}{\sqrt{Tr(V_t)}\sqrt{Tr(Q^\top QV_t)}}.\end{aligned}$$

For the extreme case where $R = 0$, the dynamics of the volatility become independent from the dynamics of the underlying what simplifies the approximation with the pivot parameter $Q^\top Q$. Indeed, the model can be rewritten as follows :

$$\begin{aligned}\frac{dS_t}{S_t} &= rdt + \sqrt{Tr(V_t)}dB_t \\ dTr(V_t) &= (\beta Tr(Q^\top Q) + 2Tr(MV_t))dt + 2\sqrt{Tr(Q^\top QV_t)}d\tilde{B}_t.\end{aligned}$$

In the dynamics of $Tr(V_t)$, the matrix $Q^\top Q$ manages the drift $(\beta Tr(Q^\top Q) + 2Tr(MV_t))$ and the volatility $2\sqrt{Tr(Q^\top QV_t)}$. However, the influence of $Q^\top Q$ in the volatility is weighted by the matrix V_t . It appears natural to consider the Cholesky matrix C of the $Q^\top Q$. The coefficient C^{12} is an orthogonal component given that it does not appears in the drift of the variance and it has a small influence in the volatility of the variance through the term $Tr(Q^\top QV_t)$. Therefore, the coefficient C^{12} has an insignificant influence on the level of the smile and will impact lightly the convexity (see Figure 2.25)

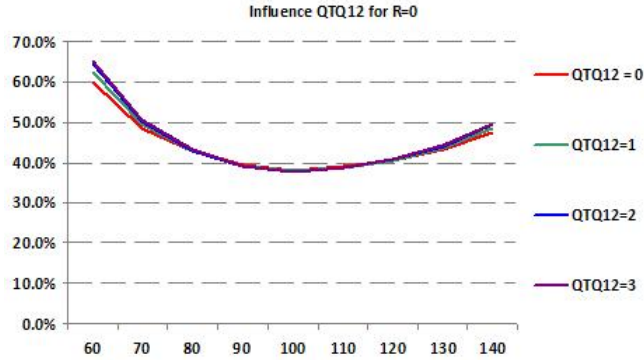


FIGURE 2.25 – Influence of C_{12} with $R = 0$

This behaviour stays the same whatever the value of the correlation. Indeed, the level of the smile for short maturities represents the level of the stochastic correlation and depends principally on $Q^\top Q$. The matrix of correlation has an influence on the volatility of the variance but the impact is small. Therefore, the component C_{12} of the Cholesky of $Q^\top Q$

has an influence only on the convexity as seen in Figure 2.26 with an example of a matrix R .

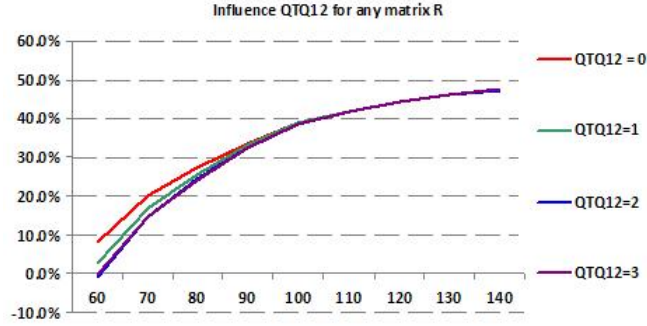


FIGURE 2.26 – Influence of C_{12} with a given matrix R

2.4.4 Influence of the initial volatility V_0

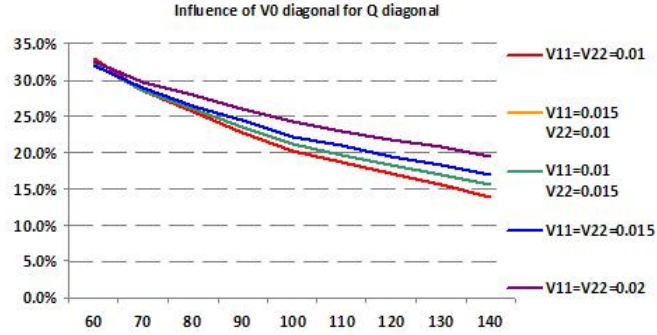
The initial volatility has a major impact on the smile. For short maturities, the smile is very close to $Tr(V_0)$ what underlines the influence of the diagonal coefficients of V_0 . Moreover, we will see the interesting impact of the non diagonal coefficients of V_0 on convexity. Let study in details the influence of the coefficients of V_0 . The set of parameters of the Wishart process is given such as V_t remains close to V_0 until the maturity $T = 6m$.

Influence of V_0^{11} and V_0^{22}

Not only the diagonal coefficients of V_0 has an influence on the level of the smile, but they also manage a little the skew and the convexity.

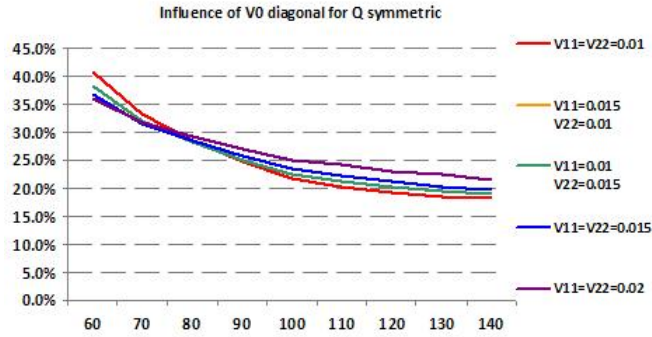
Let consider a set of parameters where V_0 and Q are diagonal matrices :

$$\beta = 4, V_0 = \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}, M = \begin{pmatrix} -0.05 & 0 \\ 0 & -0.03 \end{pmatrix}, Q = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}, R = -0.6I_n.$$


 FIGURE 2.27 – Influence of diagonal components of V_0 with diagonal matrices for Q and R

The level and the slope of the smile depend on the values of the components of V_0 (see Figure 2.27). In order to test the influence of V_0 when Q is full, we consider the new set of parameters with a symmetric matrix Q :

$$Q = \begin{pmatrix} 0.1 & 0.05 \\ 0.05 & 0.1 \end{pmatrix}.$$


 FIGURE 2.28 – Influence of diagonal components of V_0 with a symmetric matrix Q and a diagonal matrix R

As seen in Figure 2.28, the diagonal coefficients of V_0 have an influence on convexity due to the non diagonal coefficients of Q .

Influence of V_0^{12}

The matrix V_0 being symmetric, the influence of the coefficient V_0^{12} has to be studied. Moreover, the coefficient V_0^{12} can vary between two values so that the matrix V_0 remains

nonnegative. In dimension 2, the condition is given by the fact that the determinant and the trace have to be positive. The influence of V_0 is linked to the coefficients of Q and R and we will try to give intuitions about the joint influence of the parameters. We will distinguish the case where Q is diagonal and the case where Q is full.

Case where Q is diagonal

We study the influence of V_0 when Q is diagonal. The first comment is that if Q and R are diagonals, the coefficient V_0^{12} does not have any influence on the smile. Indeed, the expression of the price does not depend on V_0^{12} given that V_0^{12} does not appear in the terms $Tr(R^\top Q V_0)$, $Tr((R^\top Q)^2 V_0)$, $Tr(M V_0)$, $Tr(Q^\top Q V_0)$ and $Tr(R^\top Q V_0)^2$.

Now, we consider different set of correlation with a full matrix R . The influence of V_0^{12} depends on the non diagonal coefficients of the correlation R . When $|R_{12}|$ and $|R_{21}|$ have small values, V_0^{12} has an influence on the convexity of the smile located at high strikes (see Figure 2.29)

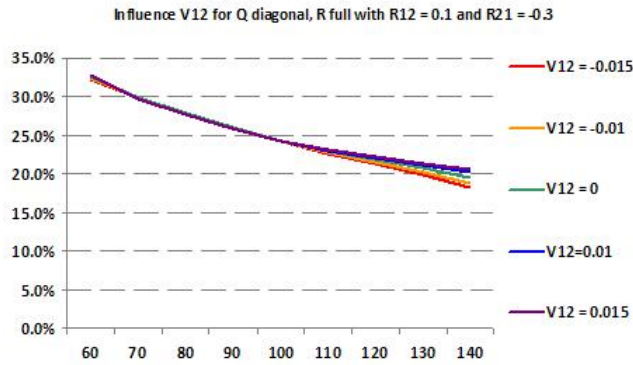


FIGURE 2.29 – Influence of V_0 with a diagonal matrix Q and a matrix R such as R_{12} and R_{21} are small

It is an interesting phenomena that the factor V_0^{12} changes the smile for high strikes letting unchanged the main part of the smile. The skew is quasi unchanged but the convexity is really impacted.

Moreover, the sign of the coefficients R_{12} and R_{21} has an influence on the convexification. In order to study this phenomena, we consider the case where $|R_{12}| = |R_{21}|$. For $R_{21} > 0$ and $R_{12} > 0$ and for $R_{21} > 0$ and $R_{12} < 0$, the convexity increases when V_0^{12} increases as shown in Figure 2.30. For $R_{21} < 0$ and $R_{12} < 0$ and for $R_{21} < 0$ and $R_{12} > 0$, Figure 2.31 underlines that the convexity decreases when V_0^{12} increases .

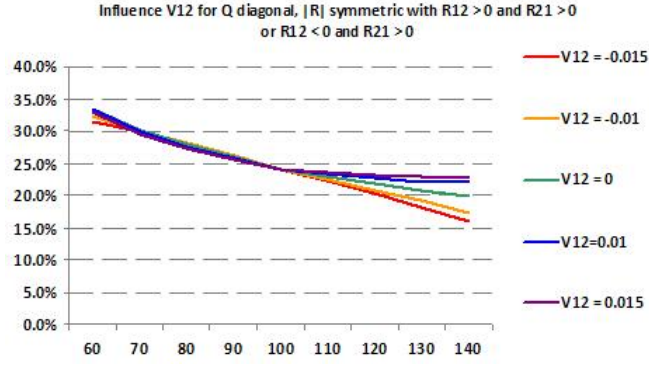


FIGURE 2.30 – Influence of V_0 with a symmetric matrix R and a diagonal matrix Q

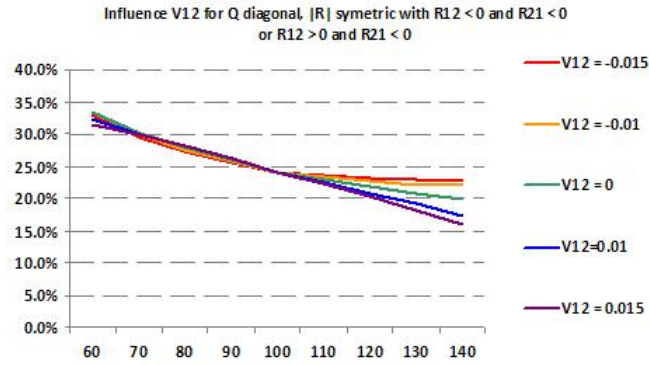


FIGURE 2.31 – Influence of V_0 with a symmetric matrix R and a diagonal matrix Q

Then, we consider high values for correlation components. For high correlations, the phenomena becomes more complex with an influence on convexity and on the slope. Indeed, we consider high non diagonal coefficients $R_{12} = -0.5$ and $R_{21} = -0.5$ and very high non diagonal coefficients $R_{12} = -0.7$ and $R_{21} = -0.7$:

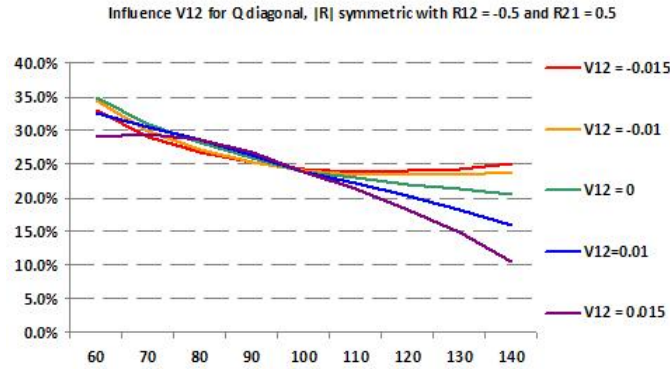


FIGURE 2.32 – Influence of V_0 with a symmetric matrix R and a diagonal matrix Q

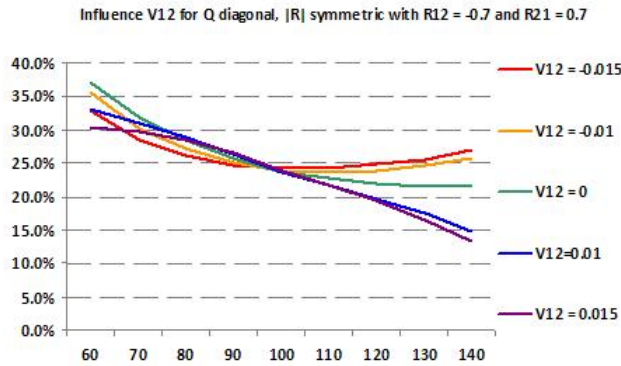


FIGURE 2.33 – Influence of V_0 with a symmetric matrix R and a diagonal matrix Q

Then, we can see that the convexity is principally located for high strikes, but there are also variations for small strikes when the non diagonal coefficients R_{12} increases.

Finally, when Q is diagonal, the influence depends on the values of the non diagonal coefficients of correlation. For small non diagonal correlation, there is a convexification for high strikes. The sign of the correlation components determines if the convexity increases or decreases when V_0^{12} increases. For high non diagonal correlations, the phenomena becomes more complex but the convexification remains principally located at high strikes with additional changes for small strikes.

Case where Q is full

We have seen that the non diagonal coefficients of Q have an strong influence on the convexity of the smile for all strikes. We will notice that the influence of V_0^{12} depends on the matrix Q and the convexity can be increased for high strikes or for small strikes. For

example, we consider the following set of parameters :

$$\beta = 4, V_0 = \begin{pmatrix} 0.02 & 0 \\ 0 & 0.02 \end{pmatrix}, M = \begin{pmatrix} -0.05 & 0 \\ 0 & -0.03 \end{pmatrix}, Q = \begin{pmatrix} 0.1 & 0.05 \\ 0.05 & 0.1 \end{pmatrix}.$$

Consider the case where the correlation is null in order to provide an intuition of the phenomena.

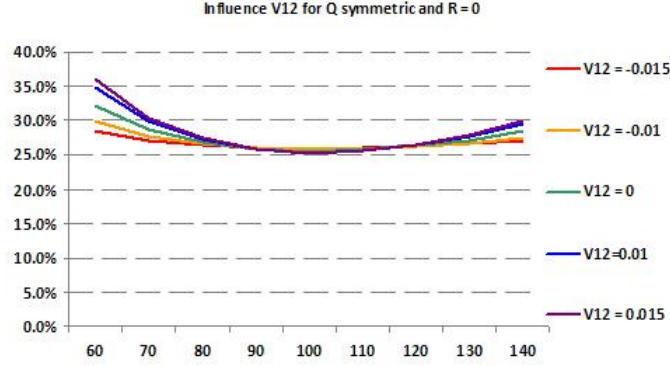


FIGURE 2.34 – Influence of V_0 with $R = 0$ and a diagonal matrix Q

When R is null, one can see that V_0^{12} has an influence on convexity for very high strikes and for very small strikes. The body of the smile is unchanged but the influence is located at the queues of the smile. Moreover, the convexity is more important for small strikes than for high strikes.

For a full matrix of correlation, we want to analyze the influence of V_0^{12} . We consider the following correlation matrix :

$$R = \begin{pmatrix} -0.6 & 0.3 \\ 0.3 & -0.6 \end{pmatrix}.$$

This correlation will change the slope of the smile but does not change the influence of V_0 as seen in Figure 2.35

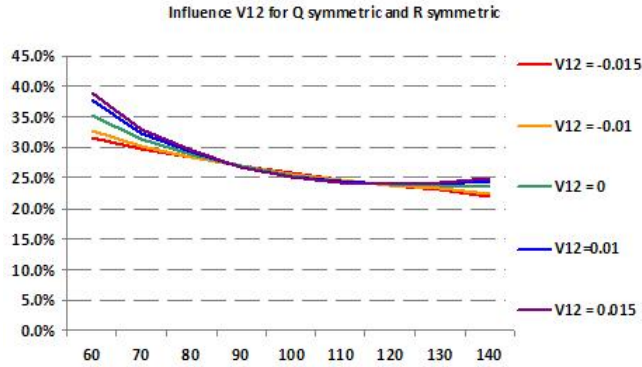


FIGURE 2.35 – Influence of V_0 with a symmetric matrix R and a symmetric matrix Q

An empirical study shows that the same phenomena remains true for a symmetric correlation. It is a really interesting observation comparing with the case of a diagonal matrix Q when the phenomena had become more complex. Moreover, for a non diagonal matrix Q , the convexification is important for small strikes whereas the convexification is important for high strikes when Q is diagonal. This behavior is more intense when the non diagonal components of Q increase.

$$Q = \begin{pmatrix} 0.1 & 0.2 \\ 0.2 & 0.1 \end{pmatrix}$$

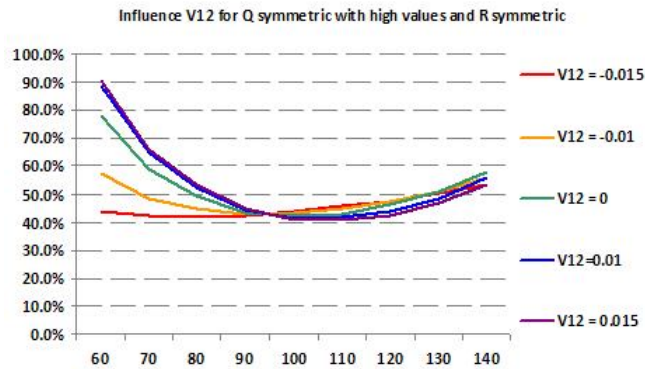


FIGURE 2.36 – Influence of V_0 with a symmetric matrix R and a symmetric matrix Q

Finally, V_0^{12} has an influence on the convexity for small strikes without changing the body of the smile for high strikes. The non diagonal coefficients of Q manage the influence of V_0 . We can summarize the influence of V_0 as follows :

- When Q is diagonal, the influence of V_0^{12} is on the convexity for high strikes.

- When the non diagonal coefficients of Q increase, the convexification progressively moves from high strikes to small strikes.
- Then, for high values of non diagonal matrices of Q , the influence corresponds to the convexity for small strikes.

We test numerically the observation we have done about the joint influence of V_0^{12} and the matrix Q . We consider the following set of parameters by considering a random matrix of correlation.

$$\beta = 4, V_0 = \begin{pmatrix} 0.02 & * \\ * & 0.02 \end{pmatrix}, M = \begin{pmatrix} -0.05 & 0 \\ 0 & -0.03 \end{pmatrix}, Q = \begin{pmatrix} 0.1 & * \\ * & 0.1 \end{pmatrix}, R = \begin{pmatrix} -0.6 & -0.4 \\ 0.2 & 0.3 \end{pmatrix}.$$

When the non diagonal coefficients of Q is null, the convexity for high strikes increases when V_0^{12} increases as shown in Figure 2.37

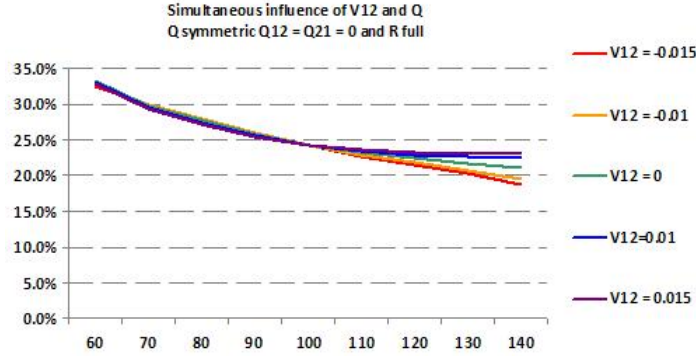


FIGURE 2.37 – Simultaneous influence of V_0 and Q with any matrix R and $Q_{12} = Q_{21} = 0$

The evolution of non diagonal components of Q has an influence of the impact of V_0^{12} . As the correlation is random, we consider a symmetric matrix Q to simplify the test but the results remains true for a general matrix Q .

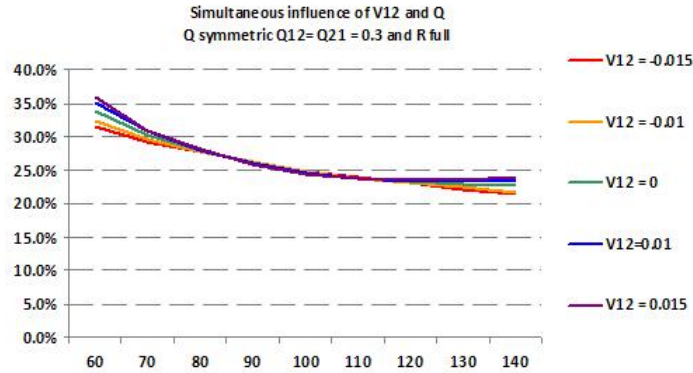


FIGURE 2.38 – Simultaneous influence of V_0 and Q with any matrix R and $Q_{12} = Q_{21} = 0.3$

When the non diagonal coefficients of Q increase, the influence on high strikes decreases and the influence for small strikes increases (see Figure 2.38)

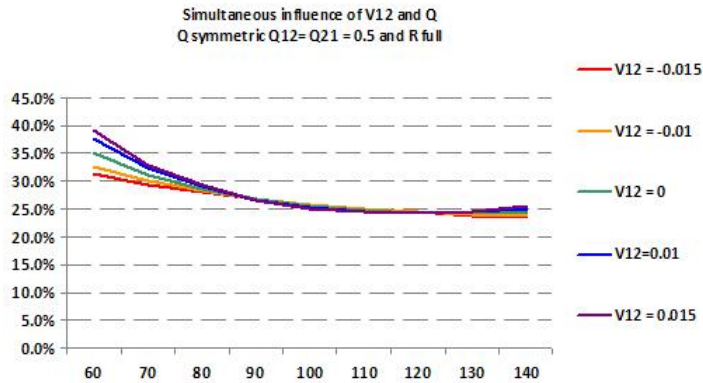


FIGURE 2.39 – Simultaneous influence of V_0 and Q with any matrix R and $Q_{12} = Q_{21} = 0.5$

Finally, when the non diagonal coefficients become important, the influence is located only on small strikes as underlined by Figure 2.38.

The influence of V_0^{12} is very interesting allowing an influence on extreme strikes. Once you have fixed some parameters that fit the level and the skew of the smile, you can choose the parameter V_0^{12} in order to fit the smile for small and high strikes. It is important to notice the joint influence of the non diagonal coefficients of Q and V_0 . Finally, the Wishart model is an efficient and useful model with a good interpretation of the influence of the parameters. The influence of the non diagonal parameters underlines the flexibility of the Wishart model with report to a standard Heston model.

2.5 Conclusion

The Wishart volatility model, which is a multifactor extension of the Heston model, underlines interesting properties. Indeed, preserving advantages of affine models provides semi-closed formula for the price of some options like Call options and Put options as well as a good understanding of the volatility behaviour through the parameters of the model. It was natural to find a procedure, based on affine properties, that allows an accurate estimation of the smile.

This paper suggests mainly some developments for smile dynamics in the Wishart volatility model. The Wishart model is flexible regarding the stochastic correlation what allows the fit of the smile dynamic. Concerning the evaluation of options, the most standard method is an evaluation by Monte Carlo methods and we present a procedure of simulation for the Wishart volatility model knowing that a standard Euler scheme is prohibited for the Wishart process. Indeed, using a relationship between a Wishart process (when the degree β is an integer) and Ornstein-Uhlenbeck processes, an extension was suggested for a general Wishart process (when a degree β is real) : in order to handle the simulation in the general case, a helpful change of probability measure was employed using the dynamics of determinant.

The link with Ornstein Uhlenbeck processes provides methods of approximation of the smile. Using an adaptation of the singular perturbations method, closed-form expressions were found concerning asymptotic smiles in the cases of short and medium maturities. Those accurate approximations can be used as a calibration tool in practice with classic conditions and give also a good idea of the parameters' effect on the smile that could be very useful for traders.

Finally, we use this efficient approximation in order to analyze the influence of the parameters. The influence is really interesting particularly with the influence of the non diagonal coefficients of the parameters.

Many points are reserved for future works : investigations about the forward smile in order to deal more complex options in the Wishart volatility model, and the generalization of this model in the interest rate world where a large number of factors is needed. One expects that this kind of model, where the volatility is a function of a Wishart process, exhibits precisely different source of risk existing in the market and provides enough flexibility for the generation of wealthy smiles.

2.6 Appendix

Appendix A : Simulation of Ornstein Uhlenbeck processes

Exact simulation of Ornstein-Uhlenbeck processes Let X be an Ornstein-Uhlenbeck process which follows the dynamics

$$dX_t = MX_t dt + Q^\top dB_t.$$

Then, it is easily shown that

$$\begin{aligned}\mathbb{E}(X_{t+h}|X_t) &= \exp(hM)X_t. \\ \text{Var}(X_{t+h}|X_t) &= \int_0^h \exp(sM)Q^\top Q \exp(sM^\top)ds.\end{aligned}$$

The function $f(t)$, defined by $f(t) = \int_0^T \exp(sM)Q^\top Q \exp(sM^\top)ds$, is solution of the differential equation

$$f'(t) = Q^\top Q + Mf(t) + f(t)M^\top.$$

Using the linearization method ([9]), the solution of this equation can be written by

$$\begin{aligned}f(t) &= g^{-1}(t)h(t). \\ (g(t) \ h(t)) &= (0 \ I_n) \exp \left[t \begin{pmatrix} M^\top & 0 \\ Q^\top Q & -M \end{pmatrix} \right].\end{aligned}$$

At last, \bar{X}_{t_i} can be deduced by $\mathcal{N}(\exp((t_i - t_{i-1})M)\bar{X}_{t_{i-1}}, f(t_i - t_{i-1}))$.

Euler scheme of Ornstein-Uhlenbeck processes Although there exists an explicit expression of \bar{X}_{t_i} with its transition law, using this method to simulate an Ornstein-Uhlenbeck process appears less effective than an Euler scheme. Indeed, as explained in the next section, the Brownian increments of the Euler scheme of the Wishart process can also be used to simulate the asset itself. Therefore, the simultaneous simulation by an Euler scheme of both the processes appears to be the most efficient method. The discretization of an Ornstein-Uhlenbeck process is given by

$$\bar{X}_{k,t_i} = (I_n + \Delta t M)\bar{X}_{k,t_{i-1}} + \sqrt{\Delta t}Q^\top \varepsilon_{k,i}, \quad 1 \leq k \leq \beta. \quad (2.25)$$

Consequently, the simulation of a Wishart process by an Euler scheme is obtained as

$$\begin{aligned}\bar{X}_{k,t_i} &= (I_n + \Delta t M)\bar{X}_{k,t_{i-1}} + \sqrt{\Delta t}Q^\top \varepsilon_{k,i}, \quad 1 \leq k \leq \beta. \\ \bar{V}_{t_i} &= \sum_{k=1}^{\beta} \bar{X}_{k,t_i} \bar{X}_{k,t_i}^\top.\end{aligned}$$

Finally, the simulation of a Wishart process in the case $\beta \in \mathbb{N}$ is easy. The case $\beta \in \mathbb{R}$ is deduced by the change of the probability measure detailed in the previous section. For

financial applications, the Wishart volatility model is a stochastic volatility model where the volatility is driven by the trace of a Wishart process. The parameters of the volatility process will be calibrated on market prices. Hence, the simulation of a Wishart process in its general form (β real deduced from calibration) is essential for an evaluation by Monte Carlo methods in the Wishart volatility model and will be developed in the next section.

Appendix B : Link between M and Q for a short maturity smile

In dimension 2, the expression of the Cholesky matrix C related to the matrix $\Sigma = Q^\top Q$ can be written in the short maturities case as :

$$\begin{aligned} C_{11} &= \sqrt{2m_1\Sigma_{11}}\sqrt{\varepsilon}. \\ C_{21} &= \frac{m_1 + m_2}{\sqrt{m_1}} \frac{\Sigma_{21}}{\sqrt{2\Sigma_{11}}} \sqrt{\varepsilon}. \\ C_{22} &= \sqrt{2\Sigma_{22}m_2 - \frac{(m_1 + m_2)^2}{m_1} \frac{(\Sigma_{21})^2}{2\Sigma_{11}}} \sqrt{\varepsilon}. \end{aligned}$$

Then, the matrix C must have the following form

$$C = \sqrt{\varepsilon} \begin{pmatrix} C_{11} & 0 \\ C_{21} & C_{22} \end{pmatrix}$$

and the matrix is found by a rotation R such as $Q = CR$. Then, we deduce that Q can be written as follows :

$$\begin{aligned} Q &= \sqrt{\varepsilon} \tilde{Q} \\ \tilde{Q} &= \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}. \end{aligned}$$

The real proof is given by a double recursive proof.

First, let us prove that the order of C_{ji} is $\sqrt{\varepsilon}$ for all $i + 1 \leq j \leq n$.

The property is true for the first term $C_{11} = \sqrt{2m_1\Sigma_{11}}\sqrt{\varepsilon}$.

Let us assume the property until the range n on the set $\mathcal{A}_n = \{C_{ji}, \forall i + 1 \leq j \leq n\}$ and we are looking to prove it on the set \mathcal{A}_{n+1} .

Let be $1 \leq i \leq n + 1$, we can write :

$$C_{n+1,i} = \frac{\Omega_{i,n+1} - \sum_{k=1}^{i-1} C_{ik}C_{n+1,k}}{C_{ii}}.$$

Then, the reader can prove easily (by a clear recursive proof or an explicit calculation) that the order of $C_{n+1,k}$ is $\sqrt{\varepsilon}$ for all $k \leq i$. Since $k \leq i$, the order of C_{ik} is also $\sqrt{\varepsilon}$ by the recursive hypothesis, one can deduce that the numerator of $C_{n+1,i}$ is at order ε . The order of the denominator is $\sqrt{\varepsilon}$ what proves the result.

Appendix C : Standard calculations in the Black-Scholes model

Recall the pricing notations in the Black-Scholes model with the log spot denoted $y = \log(x)$

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}(y + \log(\frac{e^{+r(T-t)}}{K})) + \frac{1}{2}\sigma\sqrt{T-t}.$$

$$d_0 = \frac{1}{\sigma\sqrt{T-t}}(y + \log(\frac{e^{+r(T-t)}}{K})) - \frac{1}{2}\sigma\sqrt{T-t}.$$

The following derivatives can be deduced easily

$$\frac{\partial d_1}{\partial x} = \frac{\partial d_0}{\partial x} = \frac{1}{\sigma x\sqrt{T-t}}.$$

$$\frac{\partial d_1}{\partial y} = \frac{\partial d_0}{\partial y} = \frac{1}{\sigma\sqrt{T-t}}.$$

$$\frac{\partial d_1}{\partial \sigma} = -\frac{1}{\sigma^2\sqrt{T-t}}\log(\frac{xe^{r(T-t)}}{K}) + \frac{1}{2}\sqrt{T-t}.$$

$$\frac{\partial d_0}{\partial \sigma} = -\frac{1}{\sigma^2\sqrt{T-t}}\log(\frac{xe^{r(T-t)}}{K}) - \frac{1}{2}\sqrt{T-t}.$$

Using the previous expression, the Greeks of a Call option can be calculated

$$\Delta^x = \frac{\partial C_{BS}}{\partial x} = N(d_1), \quad \Delta^y = \frac{\partial C_{BS}}{\partial y} = xN(d_1).$$

$$\Gamma^x = \frac{\partial^2 C_{BS}}{\partial x^2} = \frac{1}{\sigma x\sqrt{T-t}}N'(d_1), \quad \Gamma^y = \frac{\partial^2 C_{BS}}{\partial y^2} = x^2\Gamma^x + \Delta^y.$$

Hence, recursive calculation by considering Formula (2.20) and the Greeks (delta and gamma) yields

$$\begin{aligned} \frac{\partial^3 C_{BS}}{\partial y^3} &= \Gamma^y + (\Gamma^y - \Delta^y)\frac{(a-y)}{\sigma^2(T-t)}. \\ \frac{\partial^4 C_{BS}}{\partial y^4} &= \frac{\partial^3 C_{BS}}{\partial y^3} + (\frac{\partial^3 C_{BS}}{\partial y^3} - \frac{\partial^2 C_{BS}}{\partial y^2})\frac{(a-y)}{\sigma^2(T-t)} - \frac{1}{\sigma^2(T-t)}(\Gamma^y - \Delta^y). \\ \frac{\partial^5 C_{BS}}{\partial y^5} &= \frac{\partial^4 C_{BS}}{\partial y^4} + (\frac{\partial^4 C_{BS}}{\partial y^4} - \frac{\partial^3 C_{BS}}{\partial y^3})\frac{(a-y)}{\sigma^2(T-t)} - \frac{2}{\sigma^2(T-t)}(\frac{\partial^3 C_{BS}}{\partial y^3} - \frac{\partial^2 C_{BS}}{\partial y^2}). \\ \frac{\partial^6 C_{BS}}{\partial y^6} &= \frac{\partial^5 C_{BS}}{\partial y^5} + \frac{(a-y)}{\sigma^2(T-t)}(\frac{\partial^5 C_{BS}}{\partial y^5} - \frac{\partial^4 C_{BS}}{\partial y^4}) - \frac{3}{\sigma^2(T-t)}(\frac{\partial^4 C_{BS}}{\partial y^4} - \frac{\partial^3 C_{BS}}{\partial y^3}). \end{aligned}$$

The expression of the successive derivatives with respect to σ is given by

$$Vega^{BS} = \frac{\partial C_{BS}}{\partial \sigma} = \sigma(T-t)(\Gamma^y - \Delta^y).$$

$$Vomma^{BS} = \frac{\partial^2 C_{BS}}{\partial \sigma^2} = (T-t)d_0d_1(\Gamma^y - \Delta^y).$$

Bibliographie

- [1] Abdelkoddousse A. and A. Alfonsi (2010), “Exact and high order discretization schemes for Wishart processes and their affine extensions”, *Working paper*, [http ://hal.archives-ouvertes.fr/docs/00/49/13/71/PDF/](http://hal.archives-ouvertes.fr/docs/00/49/13/71/PDF/).
- [2] Bergomi L. (2004), “Smile Dynamics”, *Risk*, September, 117-123.
- [3] Bensusan H. (2010), “Risque de taux et de longévité : Modélisation dynamique et applications aux produits dérivés et à l’assurance-vie”, Dissertation, *Ecole polytechnique*, working paper.
- [4] Benabid A., H. Bensusan and N. El Karoui (2010), “Wishart Stochastic Volatility : Asymptotic Smile and Numerical Framework”, Submitted at *Journal of Computational Finance*.
- [5] Bru M.F. (1991), “Wishart Processes”, *Journal of Theoretical Probability*, 4, 724-743.
- [6] Carr P. and D. Madan (1999), “Option valuation using the fast Fourier transform”, *Journal of Computational Finance*, 2, 61-73.
- [7] Christoffersen P., S. Heston and K. Jacobs (2009), “The Shape and Term Structure of the Index Option Smirk : Why Multifactor Stochastic Volatility Models Work so Well”, [http ://ssrn.com/abstract=961037](http://ssrn.com/abstract=961037).
- [8] Cox J., J. Ingersoll and S. Ross (1985), “A Theory of the Term Structure of Interest Rates”, *Econometrica*, 53, 2, 385-407.
- [9] Da Fonseca J., M. Grasselli and C. Tebaldi (2005), “Wishart Multi-dimensional stochastic volatility”, *Working paper ESILV RR-31*.
- [10] Da Fonseca J., M. Grasselli and C. Tebaldi (2007), “Option pricing when correlations are stochastic : an analytical framework”, *Review of Derivatives Research*, 10, 2, 151-180.
- [11] Da Fonseca J., M. Grasselli and C. Tebaldi (2008), “A Multifactor Volatility Heston Model”, *Quantitative Finance*, 8, 6, 591-604.
- [12] Dai Q. and K.J. Singleton (2000), “Specification Analysis of Affine Term Structure Models”, *The Journal of Finance*, LV, 5, 1943-1978.
- [13] Duffie D., D. Filipovic and W. Schachermayer (2003), “Affine processes and applications in finance”, *Annals of Applied Probability*, 13, 984-1053.

- [14] Donati-Martin C., Y. Doumerc, H. Matsumoto and M. Yor (2004), “Some properties of the Wishart processes and a matrix extension of the Hartman-Watson laws”, *Publications of the Research Institute for Mathematical Sciences*, 40, 4, 1385-1412.
- [15] Durrleman V. (2004), “From Implied to Spot Volatilities”, Dissertation, *Princeton university*, [http ://www.cmap.polytechnique.fr/%7Evaldo/papers/dissertation.pdf](http://www.cmap.polytechnique.fr/%7Evaldo/papers/dissertation.pdf).
- [16] Fouque J.P., G. Papanicolaou, R. Sircar and K. Solna (2003), “Singular Perturbations in Option Pricing”, *SIAM Journal on Applied Mathematics*, 63, 5, 1648-1665.
- [17] Fouque J.P., G. Papanicolaou, R. Sircar and K. Solna (2003), “Multiscale stochastic volatility asymptotics”, *SIAM Journal Multiscale Modeling and Simulation*, 2, 1, 22-42.
- [18] Fouque J.P. and C.H. Han (2005), “Evaluation of Compound Options using Perturbation Approximation”, *Journal of Computational Finance*, 9, 1.
- [19] Gauthier P. and D. Possamai (2009), “Efficient simulation of the Wishart model”, *Working paper*, [http ://ssrn.com/abstract=1474728](http://ssrn.com/abstract=1474728).
- [20] Glasserman P. (2003), “Monte Carlo Methods in Financial Engineering”, *Springer*.
- [21] Gouriéroux C. (2006), “Wishart Processes for Stochastic Risk”, *Econometric Reviews, special issue on Stochastic Volatility*, 25, 1-41.
- [22] Gouriéroux C. and R. Sufana (2007), “Derivative Pricing with Multivariate Stochastic Volatility : Application to Credit Risk”, *Working Paper CREF 04-09*, [http ://www.yorku.ca/rsufana/dpmsv.pdf](http://www.yorku.ca/rsufana/dpmsv.pdf).
- [23] Hagan P., D. Kumar, A. Lesniewski and D. Woodward (2002), “Managing Smile Risk”, *Wilmott magazine*, July, 84-108.
- [24] Heston S. (1993), “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options”, *Review of Financial Studies*, 6, 2, 327-343.
- [25] Katori M. and H. Tanemura (2004), “Symmetry of matrix-valued stochastic processes and noncolliding diffusion particle systems”, *Journal of mathematical physics*, 45, 8, 3058-3085.
- [26] Boyarchenko N. and S. Levendorskii (2007), “On Errors and Bias of Fourier Transform Methods in Quadratic Term Structure Models”, *International Journal of Theoretical and Applied Finance*, 10, 2, 273-306.
- [27] Moler C. and C. Van Loan (2003), “Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later”, *SIAM review*, 45, 1, 3-49.
- [28] Osajima Y. (2007), “The Asymptotic Expansion formula of Implied Volatility for Dynamic SABR Model and FX Hybrid Model”, [http ://ssrn.com/abstract=965265](http://ssrn.com/abstract=965265).
- [29] Henry-Labordère P. (2005), “A General Asymptotic Implied Volatility for Stochastic Volatility Models”, [http ://ssrn.com/abstract=698601](http://ssrn.com/abstract=698601).
- [30] Saporta G. (2006), “Probabilités, analyse de données et statistiques”, Editions Technip.

Chapitre 3

Modèle de Wishart pour les taux d'intérêt

Introduction

Dans ce chapitre, nous abordons la modélisation financière de l'évolution des taux d'intérêt. Nous proposons un modèle à volatilité stochastique qui appartient à la classe des modèles affines et nous analysons les méthodes d'évaluation d'options, l'influence des paramètres du modèle et certaines de ses propriétés intrinsèques. Nous détaillons enfin la méthode de perturbations régulières appliquées à ce cadre.

Motivations

Ces dernières années, le monde des taux d'intérêt s'est complexifié avec des configurations inédites de courbe de taux et de surface de volatilité. Le monde des taux nécessite une modélisation bien plus complexe que le monde des actions dans la mesure où on essaie de capter la dynamique de la courbe de taux qui contient en théorie une infinité de facteurs. Il semblerait d'ailleurs que la courbe des taux initiale contienne de l'information et qu'elle ait un impact sur la volatilité des taux.

Afin de pouvoir capter certains de ces phénomènes, il est nécessaire d'avoir un modèle flexible avec un nombre suffisant de facteurs. L'introduction de certains modèles comme le modèle de Vasicek ou de Heath-Jarrow-Morton (HJM) ont permis de capter une bonne partie de l'information de marché. En effet, les modèles considèrent une dynamique de la courbe des taux avec des paramètres dépendant du temps afin de pouvoir se calibrer sur différents produits.

Il existe de nombreux produits standards et liquides dans le marché et les modèles se

doivent de reproduire les prix de ces produits. Tout d'abord, les swaps $S(T_0, T_N, k)$ sont des produits permettant d'échanger des intérêts à taux fixe (taux k) contre des intérêts à taux variables (taux Libor) sur un échéancier donné $[T_0, T_N]$. La durée $T_j - T_{j-1}$ séparant les dates successives de paiements est notée δ et la durée totale de l'échéancier du swap ($T_N - T_0$) est généralement appelée tenor. Le taux de swap sur un échéancier donné est défini comme le taux fixe k permettant d'annuler la valeur du swap. Les valeurs des taux de swap sur l'ensemble des échéanciers permettent de reconstituer une courbe de taux par une méthode dite de "stripping".

Les swaptions $Sw(T, T_0, T_N, K)$ correspondent aux options permettant d'entrer dans le futur à la date T dans un swap d'échéancier $[T_0, T_N]$ à un taux fixe K déterminé en début de contrat. Par conséquent, les paramètres de ce produit sont la maturité de l'option, le tenor du swap et le prix d'exercice. Ces produits sont standards et permettent de se couvrir contre des mouvements de taux d'intérêts. Enfin, les caplets $C(T, \delta, K)$ sont des options sur le taux Libor permettant d'échanger le taux Libor $L(T, T + \delta)$ contre un taux K fixé à l'avance. Les caplets correspondent à des swaptions sur un échéancier n'ayant qu'une seule date d'échange de flux dont le tenor δ est généralement petit. Ces prix d'options vanilles permettent de reconstituer une structure de volatilité des taux d'intérêt ce qui constitue une information essentielle de marché.

La volatilité implicite de ces options standard peut dépendre du prix d'exercice (strike), de la maturité et du tenor. Certains phénomènes ne sont généralement pas captés par la plupart des modèles utilisés en pratique comme par exemple la dynamique stochastique du skew des options. De plus, les traders aimeraient pouvoir caler les paramètres du modèle de manière à avoir un bon ajustement de la courbe des taux et de la surface de volatilité implicite des swaptions à la monnaie appelée surface de volatilité implicite ATM.

Ainsi, nous nous intéressons à la classe des modèles affines qui pourrait apporter des solutions à ces problèmes. Il existe un ensemble de différences fondamentales entre ces modèles et les modèles classiques. Les paramètres des modèles de type Vasicek et HJM peuvent dépendre du temps afin de suivre les mouvements de marché et ils sont calibrés tous les jours. Dans le cadre des modèles affines, les mouvements de marché sont captés par l'évolution de processus aléatoires dont la dynamique dépend de paramètres constants. Les modèles affines sont donc des modèles factoriels puisque ce sont uniquement les facteurs du modèle qui évoluent dans le temps. Cette différence majeure permet de conserver une propriété intéressante de stationnarité.

L'inconvénient majeur des modèles affines est que le nombre de facteurs doit être élevé de sorte à capter les phénomènes de marchés voulus. Cela peut rendre les modèles assez complexes et difficiles à utiliser. En revanche, l'avantage fondamental de ce type de modèle est qu'il permet d'avoir un processus de calibration "cohérent" dans la mesure où les paramètres

du modèle ne sont pas calibrés tous les jours ce qui peut d'ailleurs engendrer des opportunités d'arbitrages. Les mouvements de marché résultent uniquement de l'évolution des facteurs et non de l'évolution des paramètres. En associant certains phénomènes de marché à certains facteurs du modèle, nous pouvons interpréter certains mouvements de marchés et il devient possible de couvrir efficacement les risques sous-jacents à ces mouvements de marché.

Bien que la classe des modèles affines soit un concept théorique très intéressant, ces modèles sont assez peu utilisés en pratique. En effet, l'utilisation de ces modèles nécessite des études très poussées sur les méthodes numériques d'évaluation de produits ainsi que des analyses sur l'influence des paramètres. Les méthodes d'évaluation par Monte Carlo peuvent devenir assez compliquées étant donné le nombre important de facteurs mais il existe des méthodes d'évaluation efficaces pour les modèles affines qui utilisent la transformée de Fourier. De plus, il faut aussi étudier précisément l'influence des paramètres dont le calage est très important puisqu'ils ne sont pas sensés évoluer d'un jour à l'autre.

Présentation du chapitre

Les processus de Wishart ont été introduits par Bru [2] (1991) et ont été adaptés au monde de la finance par C. Gouriéroux et al. [15, 16]. La structure de type Wishart est une structure multidimensionnelle assez riche qui a permis d'établir des modèles à volatilité stochastique tels que le "Wishart Stochastic Volatility Model" décrit par J. Da Fonseca et al. [5, 6]. Inspiré par ces travaux innovants, nous introduisons un nouveau modèle affine décrivant la dynamique des taux d'intérêt sous la probabilité risque neutre à l'aide des processus de Wishart. Ce modèle a été initialement introduit par l'équipe de recherche de Crédit Agricole CIB encadrée par Christophe Michel, plus particulièrement les personnes qui s'intéressent aux propriétés des modèles affines (Sophie Alvarez, Anas Benabid et Victor Reutenauer). Ce modèle affine général est une extension multidimensionnelle du modèle de Heston dans le monde des taux d'intérêt pour lequel il existe des méthodes d'évaluation efficaces en grande dimension. De plus, il possède certaines propriétés intéressantes vis à vis de la corrélation stochastique du taux court qui contrôle le skew de certaines options.

Dans une première partie, nous présentons le modèle en détails et nous proposons une expression du prix des zéro-coupon dont le calcul pratique se fait par une méthode numérique de type Runge-Kutta. Nous proposons aussi des méthodes d'évaluation de produits vanilles (caplets et swaptions) qui fournissent une expression quasi explicite. L'évaluation se fait par des méthodes numériques efficaces mais nécessite tout de même une résolution de systèmes de Riccati emboîtés. Nous proposons plusieurs méthodes d'évaluation et nous utilisons en pratique la méthode d'approximation des moments qui est une méthode spécifique aux modèles affines et qui se révèle assez efficace.

Dans une deuxième partie, nous présentons une première application qui est la restriction du modèle de Wishart aux modèles de Heston en dimension 1 et 2. Nous étudions ces modèles ainsi que les méthodes d'évaluation des caplets et des swaptions. Nous analysons ensuite l'influence des paramètres sur la surface de volatilité des swaptions (ATM) qui permet aux utilisateurs d'avoir une intuition sur le comportement du modèle.

Nous décrivons dans une troisième partie les propriétés intrinsèques du modèle permettant de capter certains phénomènes de marché. Tout d'abord, nous abordons la différence entre un modèle de type Wishart et un modèle affine standard en explicitant les différentes structures de volatilité. Ensuite, nous considérons la corrélation stochastique du taux court avec sa volatilité qui contrôle au premier ordre la dynamique du skew des options sur taux court. Nous pouvons voir qu'il est possible d'exhiber un modèle de Wishart spécifique en dimension 2 permettant de capter cette corrélation avec un facteur du processus de Wishart. Le caplet étant une option sur le taux Libor qui est un taux court, ce modèle pourrait capter efficacement la dynamique du skew des caplets.

Enfin, nous proposons des méthodes d'approximation des prix de caplets lorsque l'échelle d'évolution des taux d'intérêt est longue par rapport à la maturité de l'option. Nous décrivons la méthode de perturbations régulières introduite en finance par J.P. Fouque et al. [12, 13, 14] que nous adaptons au monde complexe des taux d'intérêt. Nous appliquons cette méthode modèle de Wishart.

3.1 Présentation du modèle de Wishart sur les taux

Nous introduisons un modèle affine sur le taux court qui utilise le processus de Wishart. La forme assez générale du modèle permet une étude en dimension quelconque ce qui est particulièrement utile dans le monde des taux d'intérêt. Il est possible de choisir une répartition arbitraire des facteurs sur les taux et sur la volatilité. Nous exprimons le prix des zéro-coupon par une expression dépendant de fonctions qui sont solutions d'un système de Riccati. La résolution du système de Riccati est détaillée précisément.

3.1.1 Description du modèle

Le modèle présenté est un modèle sur le taux court r_t appartenant à la classe affine. Nous associons un processus vectoriel affine X_t et un processus matriciel de Wishart V_t de sorte que le couple (X_t, V_t) ait une dynamique affine. Le taux court est défini comme combinaison linéaire des composantes de ces processus. Le taux court, pouvant éventuellement devenir

négatif, est caractérisé par la dynamique suivante sous la probabilité risque neutre \mathbb{Q} :

$$\begin{cases} dX_t &= P(\mathfrak{N} - X_t)dt + \begin{pmatrix} Tr(\Pi_1 V_t) \\ \vdots \\ Tr(\Pi_p V_t) \end{pmatrix} dt + R\sqrt{diag(\alpha_i + Tr(\Sigma_i V_t) : 1 \leq i \leq h)}d\vec{Z}_t \\ dV_t &= (\Omega^\top \Omega + M V_t + V_t M^\top)dt + \sqrt{V_t}dW_t Q + Q^\top dW_t^\top \sqrt{V_t} \\ r_t &= \alpha + Tr[\Sigma V_t] + \delta^\top X_t \end{cases}$$

avec les notations suivantes :

- P est une matrice réelle $p \times p$
- R est une matrice réelle $p \times h$
- $(V_t)_t$ est un processus de Wishart à valeurs dans l'ensemble des matrices symétriques positives de taille $n \times n$
- Ω est une matrice réelle de taille $n \times n$.
- $(\Sigma_i)_{1 \leq i \leq h}$ sont des matrices symétriques positives (potentiellement diagonale si l'on veut utiliser seulement les composante diagonales du processus de Wishart)
- $(\Pi_i)_{1 \leq i \leq p}$ sont des matrices symétriques
- $(\alpha_i)_{1 \leq i \leq h}$ sont des nombres réels positifs
- Q est une matrice inversible de taille $n \times n$ et M est une matrice négative de taille $n \times n$
- Σ est une matrice symétrique (potentiellement positive) de taille $n \times n$
- α est un nombre réel (potentiellement positif)
- \mathfrak{N} et δ sont des vecteurs de taille $p \times 1$
- $(W_t)_t$ est une matrice $n \times n$ dont les composantes sont des mouvements Brownien scalaires indépendants sous la probabilité \mathbb{Q}
- $(Z_t)_t$ est un vecteur de taille $h \times 1$ dont les composantes sont des mouvements Browniens indépendants entre eux et indépendant de $(W_t)_t$ sous la probabilité \mathbb{Q}
- $diag(A_i : a \leq i \leq n)$ correspond à la matrice diagonale de taille $n \times n$ de coefficients diagonaux A_i

Afin d'alléger l'écriture de certaines expressions de ce document, nous introduisons dans la suite les notations suivantes :

$$D_0 = diag(\alpha_i : 1 \leq i \leq h)$$

$$D_1 = diag(\alpha_i + Tr(\Sigma_i V_t) : 1 \leq i \leq h)$$

La dynamique de la volatilité du taux court est décrite par le processus de Wishart V_t dont la valeur à l'infini dépend de $\Omega^\top \Omega$. En effet, la valeur de la moyenne du processus de Wishart à l'infini V_∞ peut être reliée à la matrice $\Omega^\top \Omega$ par :

$$-\Omega^\top \Omega = V_\infty M + M^\top V_\infty.$$

Ainsi, la matrice $\Omega^\top \Omega$ sert à calibrer la matrice de variance covariance à l'infini. Bru [2] propose une contrainte sur M et $\Omega^\top \Omega$ suffisante pour assurer l'existence du processus de

Wishart à toute date. Cette contrainte considère que la matrice M est définie négative et que $\Omega^\top \Omega = \beta Q^\top Q$ avec un paramètre β réel supérieur à $n - 1$. La dynamique du processus de Wishart ne dépend pas directement de V_∞ et reste positif. Cette spécification contraint la volatilité à l'infini à dépendre fortement de la matrice M et de la matrice Q avec le degré de liberté β . Dans ce premier travail, nous choisissons la spécification de M.F. Bru où $\Omega^\top \Omega = \beta Q^\top Q$ mais il sera possible de considérer une autre spécification assurant l'existence du processus et qui pourrait être plus flexible.

3.1.2 Évaluation du prix des zéro-coupon

Après avoir défini le modèle, nous étudions la dynamique du prix des zéro-coupon. Le prix du zéro-coupon $B(t, T)$ à la date t pour la maturité T est donné par :

$$B(t, T) = \mathbb{E}^{\mathbb{Q}} \left(e^{-\int_t^T r_s ds} \right).$$

Dans le cadre affine, le prix des zéro-coupon peut s'écrire comme l'exponentielle d'une combinaison linéaire des facteurs. Ainsi, l'expression du prix des zéro-coupon est donnée sous la forme suivante :

$$B(t, T) = \exp \left(Tr[A(t, T)V_t] + H^\top(t, T)X_t + C(t, T) \right).$$

Afin de décrire les fonctions A , H et C , on applique une formule d'Itô :

$$\begin{aligned} \frac{dB(t, T)}{B(t, T)} &= \left\{ Tr[\partial_t A(t, T)V_t] + \partial_t H^\top(t, T)X_t + \partial_t C(t, T) \right\} dt \\ &+ \left\{ \beta Tr[Q^\top Q A(t, T)] + Tr[(A(t, T)M + M^\top A(t, T))V_t] \right\} dt \\ &+ H^\top(t, T)P(\mathbb{N} - X_t)dt \\ &+ Tr \left[\left\{ \sum_{j=1}^p (H(t, T))^{j1} \Pi_j \right\} V_t \right] dt \\ &+ 2Tr[A(t, T)Q^\top Q A(t, T)]dt \\ &+ \frac{1}{2}Tr[D_0 R^\top H(t, T)H^\top(t, T)R]dt \\ &+ \frac{1}{2}Tr \left[\left\{ \sum_{j=1}^p (R^\top H(t, T)H^\top(t, T)R)^{jj} \Sigma_j \right\} V_t \right] dt \\ &+ 2Tr[\sqrt{V_t}dW_t Q A(t, T)] \\ &+ H^\top(t, T)R\sqrt{D_1}d\vec{Z}_t. \end{aligned}$$

Par absence d'opportunités d'arbitrage, le drift de la dynamique du prix des zéro-coupon est égal au taux court r_t . La dynamique du prix des zéro-coupon peut donc s'écrire sous la forme :

$$\frac{dB(t, T)}{B(t, T)} = r_t dt + \Gamma(t, T)dB_t,$$

où B_t est un mouvement brownien scalaire sous la probabilité risque neutre \mathbb{Q} . En identifiant les termes martingales des deux expressions précédentes, on en déduit :

$$\Gamma(t, T)dB_t = 2Tr[\sqrt{V_t}dW_tQA] + H^\top R\sqrt{D_1}d\vec{Z}_t$$

ce qui fournit l'expression suivante de la volatilité $\Gamma(t, T)$ du prix des zéro-coupon :

$$\Gamma(t, T)^2 = 4Tr[QA(t, T)V_tA^\top(t, T)Q^\top] + H^\top(t, T)RD_1R^\top H(t, T).$$

On remarque que la variance instantanée $\Gamma^2(t, T)$ est une combinaison linéaire des composantes du processus de Wishart V_t . Les coefficients de cette combinaison linéaire dépendent des fonctions A et H . La volatilité $\Gamma(t, T)$ dépend de T au travers de ces fonctions qui auront par conséquent de l'influence sur la structure par terme de volatilité. L'identification des parties à variations finies permet de déterminer ces fonctions. En effet, la contrainte sur le drift fournit le système de Riccati dont les fonctions A , H et C sont solutions.

$$\left\{ \begin{array}{l} \partial_t A(t, T) + A(t, T)M + M^\top A(t, T) + 2A(t, T)Q^\top QA(t, T) \\ + \left(\sum_{j=1}^p H(t, T)^{j1} \Pi_j \right) + \frac{1}{2} \left(\sum_{j=1}^h (R^\top H(t, T)H^\top(t, T)R)^{jj} \Sigma_j \right) = \Sigma \\ A(T, T) = 0. \\ \\ \partial_t H(t, T) - P^\top H(t, T) = \delta \\ H(T, T) = 0. \\ \\ \partial_t C(t, T) + \beta Tr[Q^\top QA(t, T)] + H^\top(t, T)P\aleph \\ + \frac{1}{2}Tr[D_0R^\top H(t, T)H^\top(t, T)R] = \alpha \\ C(T, T) = 0. \end{array} \right.$$

La résolution du système de Riccati est décrite dans la démonstration du théorème suivant :

Théorème 3.1. *Le prix des zéro-coupon peut s'exprimer sous la forme*

$$B(t, T) = \exp \left(Tr[A(t, T)V_t] + H^\top(t, T)X_t + C(t, T) \right), \quad (3.1)$$

avec les expressions suivantes :

$$\left\{ \begin{array}{lcl}
 H(t, T) & = & (P^\top)^{-1} \{ \exp[-P^\top(T-t)] - I_p \} \delta \\
 A(t, T) & = & [F(t, T)]^{-1} G(t, T) \\
 [\partial_t F(t, T) \quad \partial_t G(t, T)] & = & [F(t, T) \quad G(t, T)] \begin{pmatrix} M^\top & -\Theta_A(t) \\ 2Q^\top Q & -M \end{pmatrix} \\
 [F(T, T) \quad G(T, T)] & = & [I_n \quad 0] \\
 \Theta_A(t) & = & \Sigma - \left(\sum_{j=1}^p H(t, T)^{j1} \Pi_j \right) \\
 & & - \frac{1}{2} \left(\sum_{j=1}^h (R^\top H(t, T) H^\top(t, T) R)^{jj} \Sigma_j \right) \\
 C(t, T) & = & -\frac{\beta}{2} \{ (T-t) \text{Tr}[M] + \log(\det(F(t, T))) \} \\
 & & - ((T-t)\delta^\top - H^\top(t, T)) \aleph \\
 & & + \frac{1}{2} \text{Tr} \left[D_0 \int_t^T R^\top H(s, T) H^\top(s, T) R ds \right] \\
 & & - (T-t)\alpha \\
 P^\top \left[\int_t^T H(s, T) H^\top(s, T) ds \right] P & = & \int_t^T \exp[-P^\top(T-s)] \delta \delta^\top \exp[-P(T-s)] ds \\
 & & + (P^\top)^{-1} \{ \exp[-P^\top(T-t)] - I_p \} \delta \delta^\top \\
 & & + \delta \delta^\top \{ \exp[-P(T-t)] - I_p \} P^{-1} \\
 & & + (T-t) \delta \delta^\top \\
 \int_t^T \exp[-P^\top(T-s)] \delta \delta^\top \exp[-P(T-s)] ds & = & \exp[-P^\top(T-t)] [U(t)]^{-1} N(t) \exp[-P(T-t)] \\
 [U(t) \quad N(t)] & = & [I_n \quad 0] \exp \left\{ \begin{pmatrix} -(T-t)P^\top & (T-t)\delta \delta^\top \\ 0 & (T-t)P \end{pmatrix} \right\}
 \end{array} \right.$$

Démonstration 3.1. Afin de démontrer ce théorème, il faut résoudre le système de Riccati vérifié par les fonctions A , H et C . Tout d'abord, on peut réécrire la première équation de Riccati sous la forme suivante :

$$\partial_t A(t, T) + A(t, T)M + M^\top A(t, T) + 2A(t, T)Q^\top Q A(t, T) = \Theta_A(t),$$

en ayant défini au préalable la fonction $\Theta_A(t)$ par :

$$\Theta_A(t) = \Sigma - \left(\sum_{j=1}^p H(t, T)^{j1} \Pi_j \right) - \frac{1}{2} \left(\sum_{j=1}^h (R^\top H(t, T) H^\top(t, T) R)^{jj} \Sigma_j \right).$$

Le fait que $\Theta_A(t)$ dépende du temps empêche l'obtention d'une forme fermée à l'aide de l'exponentielle matricielle. Cependant, comme il est fait dans l'annexe A.1, on peut se ramener à une forme semi-fermée qu'il faudra ensuite résoudre numériquement par des

méthodes de type Runge-Kutta. La solution peut s'écrire sous la forme suivante :

$$\begin{cases} A(t, T) &= [F(t, T)]^{-1}G(t, T) \\ [\partial_t F(t, T) \quad \partial_t G(t, T)] &= [F(t, T) \quad G(t, T)] \begin{pmatrix} M^\top & -\Theta_A(t) \\ Q^\top Q & -M \end{pmatrix} \\ [F(T, T) \quad G(T, T)] &= [I_n \quad 0] \end{cases}$$

La résolution de la deuxième équation est immédiate :

$$H(t, T) = (P^\top)^{-1} \left\{ \exp[-P^\top(T-t)] - I_p \right\} \delta.$$

En intégrant la dernière équation, on peut écrire :

$$\begin{aligned} C(t, T) &= \beta Tr \left[\int_t^T Q^\top Q A(s, T) ds \right] + \int_t^T H^\top(s, T) P \delta ds \\ &+ \frac{1}{2} Tr \left[D_0 \int_t^T R^\top H(s, T) H^\top(s, T) R ds \right] - (T-t)\alpha. \end{aligned}$$

Calculons à présent les trois intégrales présentes dans la formule précédente :

1. Calcul de l'intégrale $\int_t^T Q^\top Q A(s, T) ds$:

La méthode de linéarisation (voir annexe A.1) appliquée à la première équation fournissait l'équation suivante :

$$-[F(t, T)]^{-1} \partial_t F(t, T) + M^\top + 2A(t, T)Q^\top Q = 0.$$

Par conséquent, le calcul de la première intégrale s'écrit sous la forme :

$$2Tr \left[\int_t^T Q^\top Q A(s, T) ds \right] = -(T-t)Tr[M] - \log(\det(F(t, T))).$$

2. Le calcul de l'intégrale $\int_t^T H^\top(s, T) P \delta ds$ se fait en utilisant la deuxième équation de Riccati. En effet, un calcul immédiat fournit :

$$\int_t^T H(s, T) P ds = -(T-t)\delta^\top - H^\top(t, T).$$

3. Calcul de l'intégrale $\int_t^T R^\top H(s, T) H^\top(s, T) R ds$:

Pour cette troisième intégrale, on utilise une méthode de linéarisation de Riccati. On peut calculer $\int_t^T H(s, T) H^\top(s, T) ds$ en utilisant :

$$\begin{aligned} P^\top \left[\int_t^T H(s, T) H^\top(s, T) ds \right] P &= \int_t^T \exp[-P^\top(T-s)] \delta \delta^\top \exp[-P(T-s)] ds \\ &+ (P^\top)^{-1} \left\{ \exp[-P^\top(T-t)] - I_p \right\} \delta \delta^\top \\ &+ \delta \delta^\top \left\{ \exp[-P(T-t)] - I_p \right\} P^{-1} \\ &+ (T-t)\delta \delta^\top. \end{aligned}$$

En utilisant l'annexe A.2, on déduit l'expression de l'intégrale suivante

$\int_t^T \exp[-P^\top(T-s)]\delta\delta^\top \exp[-P(T-s)]ds$ sous la forme :

$$\int_t^T \exp[-P^\top(T-s)]\delta\delta^\top \exp[-P(T-s)]ds = \exp[-P^\top(T-t)][U(t)]^{-1}N(t) \exp[-P(T-t)]$$

$$\text{où } [U(t) \ N(t)] = [I_n \ 0] \exp \left\{ \begin{pmatrix} -(T-t)P^\top & (T-t)\delta\delta^\top \\ 0 & (T-t)P \end{pmatrix} \right\}.$$

Ainsi, on obtient une formule quasi-fermée pour le prix des zéro-coupon. La résolution du système se fait en utilisant un schéma numérique de Runge-Kutta à l'ordre 4.

3.1.3 Description de la probabilité forward neutre

La probabilité U-forward neutre est définie de la manière suivante :

$$\frac{d\mathbb{Q}_U}{d\mathbb{Q}_{\mathcal{F}_t}} = \frac{\exp\left(-\int_t^U r_s ds\right)}{B(t, U)}.$$

La dynamique du prix des zéro-coupon peut s'écrire :

$$\frac{dB(t, U)}{B(t, U)} = r_t dt + 2Tr[\sqrt{V_t}dW_t Q A(t, U)] + H^\top(t, U)R\sqrt{D_1}d\vec{Z}_t.$$

Ainsi, le théorème de Girsanov fournit l'expression des nouveaux mouvements browniens sous la probabilité U-forward neutre :

$$dW_t^U = dW_t - 2\sqrt{V_t}A(t, U)Q^\top dt$$

$$d\vec{Z}_t^U = d\vec{Z}_t - \sqrt{D_1}R^\top H(t, U)dt.$$

Posons $M'(t, U) = M + 2Q^\top Q A(t, U)$. La dynamique du modèle sous la probabilité U-forward peut se réécrire comme suit :

$$\begin{cases} dX_t = P(\mathbb{N} - X_t)dt + \begin{pmatrix} Tr(\Pi_1 V_t) \\ \vdots \\ Tr(\Pi_p V_t) \end{pmatrix} dt + RD_1 R^\top H(t, U)dt + R\sqrt{D_1}d\vec{Z}_t^U \\ dV_t = \left(\beta Q^\top Q + M'(t, U)V_t + V_t(M'(t, U))^\top \right) dt \\ \quad + \sqrt{V_t}dW_t^U Q + Q^\top (dW_t^U)^\top \sqrt{V_t} \\ r_t = \alpha + Tr[\Sigma V_t] + \delta^\top X_t. \end{cases}$$

3.1.4 Transformation de Laplace des facteurs

Dans un modèle affine, l'expression de la transformée de Laplace va être au centre de nombreux développements dont l'évaluation de produits dérivés. Nous définissons la transformée de Laplace des facteurs sous une probabilité forward neutre \mathbb{Q}_U de la manière suivante :

$$\Phi(t, T, U, \Gamma, \Lambda, V_t, X_t) = \mathbb{E}_t^{\mathbb{Q}_U} \left[\exp \left(Tr[\Gamma V_T] + \Lambda^\top X_T \right) \right],$$

où Γ est une matrice de taille $n \times n$ et Λ un vecteur de taille p .

Le modèle appartient au cadre affine, nous allons donc chercher la solution sous la forme d'une exponentielle d'une combinaison linéaire des facteurs :

$$\Phi(t, T, U, \Gamma, \Lambda, V_t, X_t) = \exp \left(\text{Tr}[D(t, T, U, \Gamma, \Lambda)V_t] + [J(t, T, U, \Gamma, \Lambda)]^\top X_t + Z(t, T, U, \Gamma, \Lambda) \right).$$

Un développement stochastique à l'aide de la formule d'Itô, en exploitant le fait que $\Phi(t, T, U, \Gamma, \Lambda, V_t, X_t)$ soit une martingale, fournit un nouveau système de Riccati qui s'écrit :

$$\left\{ \begin{array}{l} \partial_t D + D[M'(t, U)] + [M'(t, U)]^\top D + 2D[Q^\top Q]D \\ + \left(\sum_{j=1}^h [(R^\top H(t, U)JR)^{jj} + \frac{1}{2}(R^\top JJ^\top R)^{jj}\Sigma_j] \right) \\ + \left(\sum_{j=1}^p J(t, T, U, \Gamma, \Lambda)^{j1}\Pi_j \right) = 0 \\ D(T, T, U, \Gamma, \Lambda) = \Gamma. \\ \\ \partial_t J - P^\top J = 0 \\ J(T, T, U, \Gamma, \Lambda) = \Lambda. \\ \\ \partial_t Z(t, T) + \beta \text{Tr}[Q^\top QD] + J^\top P\aleph + \text{Tr}[D_0 \{ \frac{1}{2}R^\top JJ^\top R + R^\top H(t, U)J^\top R \}] = 0 \\ Z(T, T, U, \Gamma, \Lambda) = 0. \end{array} \right.$$

Ce système de Riccati fait intervenir la fonction H apparaissant dans l'expression du prix des zéro-coupon. Bien que les paramètres dépendent du temps, ce système peut être résolu quasi- explicitement comme l'énonce le théorème suivant.

Théorème 3.2. *La transformée de Laplace $\Phi(t, T, U, \Gamma, \Lambda, V_t, X_t)$ s'exprime de la manière suivante :*

$$\Phi(t, T, U, \Gamma, \Lambda, V_t, X_t) = \exp \left(\text{Tr}[D(t, T, U, \Gamma, \Lambda)V_t] + [J(t, T, U, \Gamma, \Lambda)]^\top (t, T)X_t + Z(t, T, U, \Gamma, \Lambda) \right),$$

avec les expressions suivantes :

$$\left\{ \begin{array}{lcl}
 J(t, T, U, \Gamma, \Lambda) & = & \exp[-P^\top(T-t)]\Lambda \\
 D(t, T, U, \Gamma, \Lambda) & = & [F(t, T)]^{-1}G(t, T) \\
 \begin{array}{l} [\partial_t F(t, T) \quad \partial_t G(t, T)] \\ [F(T, T) \quad G(T, T)] \end{array} & = & \begin{array}{l} [F(t, T) \quad G(t, T)]\Pi(t, U) \\ [I_n \quad \Gamma] \end{array} \\
 \Pi(t, U) & = & \begin{pmatrix} [M'(t, U)]^\top & \Theta_D(t) \\ 2Q^\top Q & -M'(t, U) \end{pmatrix} \\
 \Theta_D(s) & = & \begin{array}{l} \frac{1}{2} \sum_{j=1}^h (R^\top J(s, T, U, \Gamma, \Lambda) J^\top(s, T, U, \Gamma, \Lambda) R)^{jj} \Sigma_j \\ + \sum_{j=1}^h (R^\top H(s, U) J^\top(s, T, U, \Gamma, \Lambda) R)^{jj} \Sigma_j \\ + \sum_{j=1}^p J(s, T, U, \Gamma, \Lambda)^{j1} \Pi_j \end{array} \\
 Z(t, T, U, \Gamma, \Lambda) & = & \begin{array}{l} + \frac{\beta}{2} [\log(\det(F(t, U))) - \log(\det(F(T, U)))] \\ - \frac{\beta}{2} \log(\det(F(t, T, U, \Gamma, \Lambda))) \\ + (\Lambda^\top - J^\top(t, T, U, \Gamma, \Lambda)) \aleph \\ + \frac{1}{2} Tr \left[D_0 \int_t^T R^\top J(s, T, U, \Gamma, \Lambda) J^\top(s, T, U, \Gamma, \Lambda) R ds \right] \\ - Tr \left[D_0 \int_t^T R^\top H(s, U) J^\top(s, T, U, \Gamma, \Lambda) R ds \right] \end{array} \\
 \int_t^T J(s, T, U, \Gamma, \Lambda) J^\top(s, T, U, \Gamma, \Lambda) ds & = & \exp[-P^\top(T-t)][U(t)]^{-1}N(t) \exp[-P(T-t)] \\
 [U(t) \quad N(t)] & = & [I_n \quad 0] \exp \left\{ \begin{pmatrix} -(T-t)P^\top & (T-t)\Lambda\Lambda^\top \\ 0 & (T-t)P \end{pmatrix} \right\} \\
 P^\top \int_t^T H(s, U) J^\top(s, T, U, \Gamma, \Lambda) ds & = & \begin{array}{l} \exp[-P^\top(U-t)][X(t)]^{-1}Y(t) \exp[-P(T-t)] \\ - \delta\Lambda^\top [I_p - \exp[-P(T-t)]]P^{-1} \end{array} \\
 [X(t) \quad Y(t)] & = & [I_n \quad 0] \exp \left\{ \begin{pmatrix} -(T-t)P^\top & (T-t)\delta\Lambda^\top \\ 0 & (T-t)P \end{pmatrix} \right\}.
 \end{array} \right.$$

Démonstration 3.2. Pour démontrer ce théorème, il faut résoudre le système de Riccati vérifié par les fonctions intervenant de la transformée de Laplace. La première équation se résout comme précédemment avec une méthode de linéarisation. Concernant la deuxième équation, la résolution est immédiate et fournit :

$$J(t, T, U, \Gamma, \Lambda) = \exp(-(T-t)P^\top)\Lambda.$$

Il reste à exprimer la forme de la solution de la fonction Z . En intégrant la dernière équation,

on obtient :

$$\begin{aligned}
 Z(t, T, U, \gamma, \Lambda) &= \beta Tr \left[\int_t^T Q^\top Q D(s, T, U, \Gamma, \Lambda) ds \right] + \int_t^T J^\top(s, T, U, \Gamma, \Lambda) P \aleph ds \\
 &+ Tr \left[\frac{1}{2} D_0 \int_t^T R^\top J(s, T, U, \Gamma, \Lambda) J^\top(s, T, U, \Gamma, \Lambda) R ds \right] \\
 &+ Tr \left[\frac{1}{2} D_0 \int_t^T R^\top H(s, U) J^\top(s, T, U, \Gamma, \Lambda) R ds \right].
 \end{aligned}$$

On peut simplifier cette expression en utilisant l'équation donnée par la linéarisation de la fonction D qui mène à :

$$\begin{aligned}
 2Tr \left[\int_t^T Q^\top Q D(s, T, U, \Gamma, \Lambda) ds \right] &= -(T-t)Tr(M) - 2Tr \left[\int_t^T Q^\top Q A(s, U) ds \right] \\
 &- \log(\det(F(t, T, U, \Gamma, \Lambda))) \\
 &= \log \left(\frac{\det(F(t, U))}{\det(F(T, U))} \right) - \log(\det(F(t, T, U, \Gamma, \Lambda))).
 \end{aligned}$$

De plus, en intégrant la deuxième équation de Riccati, la deuxième intégrale se calcule par :

$$\int_t^T J^\top(s, T, U, \Gamma, \Lambda) P \aleph ds = [\Lambda^\top - J^\top(t, T, U, \Gamma, \Lambda)] \aleph.$$

Concernant les deux derniers termes, ils se calculent en utilisant l'annexe A.2

$$\left[\int_t^T J(s, T, U, \Gamma, \Lambda) J^\top(s, T, U, \Gamma, \Lambda) ds \right] = \int_t^T \exp[-P^\top(T-s)] \Lambda \Lambda^\top \exp[-P(T-s)] ds.$$

En utilisant l'annexe A.2, on peut déduire l'expression de l'intégrale

$\int_t^T \exp[-P^\top(T-s)] \Lambda \Lambda^\top \exp[-P(T-s)] ds$ sous la forme suivante :

$$\begin{aligned}
 \int_t^T \exp[-P^\top(T-s)] \Lambda \Lambda^\top \exp[-P(T-s)] ds &= \exp[-P^\top(T-t)] [U(t)]^{-1} N(t) \exp[-P(T-t)] \\
 \text{où } [U(t) \ N(t)] &= [I_p \ 0] \exp \left\{ \begin{pmatrix} -(T-t)P^\top & (T-t)\Lambda\Lambda^\top \\ 0 & (T-t)P \end{pmatrix} \right\}.
 \end{aligned}$$

Pour le dernier terme, les calculs se font par la même méthode et mènent à :

$$\begin{aligned}
 \left[P^\top \int_t^T H(s, U) J^\top(s, T, U, \Gamma, \Lambda) ds \right] &= \int_t^T \exp[-P^\top(U-s)] \delta \Lambda^\top \exp[-P(T-s)] ds \\
 &- \int_t^T [\delta \Lambda^\top \exp(-P(T-t))] \\
 &= \exp[-P^\top(U-t)] [X(t)]^{-1} Y(t) \exp[-P(T-t)] \\
 &- \delta \Lambda^\top [I_p - \exp(-P(T-t))] P^{-1},
 \end{aligned}$$

où les fonctions X et Y vérifient :

$$[X(t) \ Y(t)] = [I_p \ 0] \exp \left\{ \begin{pmatrix} -(T-t)P^\top & (T-t)\delta\Lambda^\top \\ 0 & (T-t)P \end{pmatrix} \right\}.$$

Ce théorème fournit le calcul quasi explicite de la transformée de Laplace sous une probabilité forward neutre. Cette expression permet de trouver des méthodes d'évaluation efficaces pour les options vanilles.

3.2 Méthodes d'évaluation des options vanilles

Dans le marché, certains instruments financiers sont utilisés pour calibrer les modèles des taux d'intérêt. En effet, la plupart des modèles de taux d'intérêt sont calibrés sur les prix des caplets et des swaptions. Par exemple, pour un modèle Vasicek standard avec des paramètres dépendant du temps, les options vanilles permettent de caler ces paramètres : le paramètre $\theta(t)$ est généralement calibré sur les prix de caplets et le paramètre $\lambda(t)$ sur l'antidiagonale des prix de swaptions jusqu'à la maturité T_{max} correspondant généralement aux swaptions de maturité $T = i$ années et de tenor $U = T_{max} - i$ années. De manière générale, le prix des caplets fournit de l'information sur les niveaux de volatilité des taux Libor forward et la matrice de prix des swaptions fournit de l'information sur la corrélation entre les taux Libor forward successifs.

Dans le modèle de Wishart, il faut étudier les méthodes d'évaluation des options vanilles. Ainsi, nous décrivons les expressions quasi explicites des prix de caplets et des prix de swaptions dont l'évaluation se fait par des méthodes numériques rapides.

3.2.1 Évaluation des swaptions

Dans cette partie nous abordons l'évaluation des swaptions qui peut se faire par FFT ou par la méthode des moments. L'évaluation ne peut pas se faire directement et il faut mettre en œuvre une approximation. En effet, le taux de swap forward fixé en T pour un swap commençant en T_0 et se terminant en T_N avec une fréquence de paiement de δ est défini par :

$$S_T(T_0, T_N, \delta) = \frac{B(T, T_0) - B(T, T_N)}{\sum_{i=1}^N \delta B(T, T_i)}.$$

Une swaption est une option d'entrée à la maturité T dans un swap d'échéancier $[T_0, T_N]$ au prix d'exercice K fixé à l'avance. Par conséquent, son prix est donné par la formule suivante :

$$\begin{aligned} Sw(T, T_0, T_N, \delta, K) &= B(0, T) \mathbb{E}^{Q_T} \left[(S_T(T_0, T_N, \delta) - K)^+ \sum_{i=1}^N \delta B(T, T_i) \right] \\ &= B(0, T) \mathbb{E}^{Q_T} \left[\left(B(T, T_0) - B(T, T_N) - K \sum_{i=1}^N \delta B(T, T_i) \right)^+ \right]. \end{aligned}$$

Approximation dans un modèle affine

Il est impossible de trouver une formule fermée à ces prix et il est nécessaire de faire une approximation. L'approximation classique pour les modèles affines est de supposer que $Y_T = B(T, T_N) + K \sum_{i=1}^N \delta B(T, T_i)$ s'écrit sous la forme d'une exponentielle affine. L'idée est de poser :

$$H_s = \frac{B(s, T_N) + K \sum_{i=1}^N \delta B(s, T_i)}{B(s, T)}, \quad s \leq T.$$

Ce processus H_s est martingale sous \mathbb{Q}_T . En effet, sa dynamique peut s'écrire :

$$\begin{aligned} \frac{dH_s}{H_s} &= 2Tr \left[\sqrt{V_s} dW_s^T Q \left\{ \frac{B(s, T_N)A(s, T_N) + K \sum_{i=1}^N B(s, T_i)A(s, T_i)}{B(s, T_N) + K \sum_{i=1}^N B(s, T_i)} - A(s, T) \right\} \right] \\ &+ \left\{ \frac{B(s, T_N)H^\top(s, T_N) + K \sum_{i=1}^N B(s, T_i)H^\top(s, T_i)}{B(s, T_N) + K \sum_{i=1}^N B(s, T_i)} - H^\top(s, T) \right\} R\sqrt{D_1} dB_s^T. \end{aligned}$$

La première approximation consiste à considérer que

$$\frac{dH_s}{H_s} \sim 2Tr \left[\sqrt{V_s} dW_s^T Q \bar{A}(s) \right] + \bar{H}^\top(s, T) R\sqrt{D_1} dB_s^T,$$

avec

$$\begin{aligned} \bar{A}(s) &= \frac{B(0, T_N)A(s, T_N) + K \sum_{i=1}^N B(0, T_i)A(s, T_i)}{B(0, T_N) + K \sum_{i=1}^N B(0, T_i)} - A(s, T), \\ \bar{H}(s) &= \frac{B(0, T_N)H(s, T_N) + K \sum_{i=1}^N B(0, T_i)H(s, T_i)}{B(0, T_N) + K \sum_{i=1}^N B(0, T_i)} - H(s, T). \end{aligned}$$

Cette approximation est assez classique concernant les swaptions ce qui revient à supposer que les poids ont une volatilité très faible. La deuxième approximation naturelle consiste à poser que :

$$H_T \sim \exp \left[Tr(\bar{A}(T)V_T) + \bar{H}^\top(T)X_T + \bar{C}(T) \right].$$

Pour calculer $\bar{C}(T)$, on utilise la propriété de martingale de H_T à l'ordre 0 qui permet d'écrire $\mathbb{E}^{Q_T}[H_T] = H_0$. Cela fournit :

$$\bar{C}(T) = \log(H_0) - \log(\Phi(0, T, T, \bar{A}(T), \bar{H}^\top(T), V_0, X_0)).$$

Finalement, on peut écrire l'approximation de Y_T sous la forme d'une exponentielle affine :

$$Y_T = B(T, T_N) + K \sum_{i=1}^N \delta B(T, T_i) \sim \exp \left[Tr(\bar{A}(T)V_T) + \bar{H}^\top(T)X_T + \bar{C}(T) \right].$$

Évaluation des swaptions par la méthode des moments

Connaissant la transformée de Laplace des facteurs et ayant fait l'approximation sur la variable $Y_T = B(T, T_N) + K \sum_{i=1}^N \delta B(T, T_i)$, il est facile de calculer les moments de Y_T . En effet, la forme exponentielle trouvée fournit un moyen facile de calculer tous les moments d'ordre j :

$$\begin{aligned} m_j = \mathbb{E}(Y_T^j) &= \mathbb{E}^{Q_T} \left[\exp \left[Tr(j\bar{A}(T)V_T) + j\bar{H}^\top(T)X_T + j\bar{C}(T) \right] \right] \\ &= e^{j\bar{C}(T)} \Phi(0, T, T, j\bar{A}(T), j\bar{H}(T), V_0, X_0). \end{aligned}$$

En considérant les moments de cette variable, on peut utiliser des méthodes de développements de type Edgeworth ou Gram Charlier. Ces méthodes reconstituent des valeurs d'options en utilisant un développement selon les moments (ou cumulants) jusqu'à un certain ordre. On détaille le calcul des prix d'options par la méthode des moments.

Soit Y une variable aléatoire. La transformée de Fourier $\Phi^Y(k) = \mathbb{E}(e^{ikY})$ de la variable Y peut être définie à partir des cumulants de la +manière suivante :

$$\ln[\Phi^Y(k)] = \sum_{j=1}^{\infty} c_j \frac{(ik)^j}{j!}.$$

De manière standard, les cumulants c_j de la variable Y se calculent facilement à partir de ses moments m_j :

$$\begin{aligned} c_1 &= m_1, & c_2 &= m_2 - m_1^2, & c_3 &= m_3 - 3m_1m_2 + 2m_1^3, \\ c_4 &= m_4 - 4m_1m_3 + 3m_2^2 + 12m_1^2m_2 - 6m_1^4, \\ c_5 &= m_5 - 5m_1m_4 - 10m_2m_3 + 20m_1^2m_3 + 30m_1m_2^2 - 60m_1^3 + 24m_1^5, \\ c_6 &= m_6 - 6m_1m_5 - 15m_2m_4 + 30m_1^2m_4 - 10m_3^2 + 120m_1m_2m_3 - 120m_1^3m_3 \\ &\quad + 30m_2^3 - 270m_1^2m_2^2 + 360m_1^4m_2 - 120m_1^6, \\ c_7 &= m_7 - 7m_1m_6 - 21m_2m_5 - 35m_3m_4 + 140m_1m_3^2 - 630m_1m_2^3 + 210m_1m_2m_4 \\ &\quad - 1260m_1^2m_2m_3 + 42m_1^2m_5 + 2520m_1^3m_2^2 - 210m_1^3m_2^2 + 210m_2^2m_3 + 840m_1^4m_3 \\ &\quad - 2520m_1^5m_2 - 720m_1^7. \end{aligned}$$

Connaissant les cumulants, on peut faire un développement de la densité de la variable Y notée $f^Y(y)$

$$f^Y(y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iky} \Phi^Y(k) dk = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iky} e^{\sum_{j=1}^{\infty} c_j \frac{(ik)^j}{j!}} dk = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iky} e^{ikc_1 - \frac{k^2}{2}c_2} e^{\Lambda} dk,$$

où $e^\Lambda = e^{\sum_{j=3}^{\infty} c_j \frac{(ik)^j}{j!}}$ est approximé au deuxième ordre en ne considérant que les 7 premiers cumulants du développement :

$$\begin{aligned} e^\Lambda &\sim 1 + \Lambda + \frac{\Lambda^2}{2} \\ &\sim 1 - ik^3 \left(\frac{c_3}{3!} \right) + k^4 \left(\frac{c_4}{4!} \right) + ik^5 \left(\frac{c_5}{5!} \right) - k^6 \left(\frac{c_6}{6!} + \frac{1}{2} \left(\frac{c_3}{3!} \right)^2 \right) - ik^7 \left(\frac{c_7}{7!} + \frac{c_3}{3!} \frac{c_4}{4!} \right). \end{aligned}$$

Ainsi, on peut écrire le développement de f^Y sous la forme suivante :

$$f(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iky} e^{ikc_1 - \frac{k^2}{2}c_2} \left(1 + \sum_{i=3}^7 \theta_n k^n \right) dk$$

Enfin, il existe plusieurs méthodes d'approximation (Edgeworth, Gram Charlier) permettant de calculer les prix d'options. On rappelle l'expression des polynômes de Hermite jusqu'à l'ordre 7 :

$$\begin{aligned} H_0 &= 1, & H_1 &= x, & H_2 &= x^2 - 1, & H_3 &= x^3 - 3x, & H_4 &= x^4 - 6x^2 + 3, \\ H_5 &= x^5 - 10x^3 + 15x, & H_6 &= x^6 - 15x^4 + 45x^2 - 15, & H_7 &= x^7 - 21x^5 + 105x^3 + 105x. \end{aligned}$$

La méthode de Gram Charlier [4] consiste en une régression sur les polynômes d'Hermite sous la forme

$$\begin{aligned} \mathbb{E}[1_{Y>a}] &= \mathcal{N}\left(\frac{c_1 - a}{\sqrt{c_2}}\right) + \sum_{k=3}^{\infty} (-1)^{k-1} q_k H_{k-1}\left(\frac{c_1 - a}{\sqrt{c_2}}\right) \Phi\left(\frac{c_1 - a}{\sqrt{c_2}}\right). \\ \mathbb{E}[Y 1_{Y>a}] &= c_1 \mathcal{N}\left(\frac{c_1 - a}{\sqrt{c_2}}\right) + \sqrt{c_2} \Phi\left(\frac{c_1 - a}{\sqrt{c_2}}\right) \\ &\quad + \sum_{k=3}^{\infty} (-1)^k q_k \left(-a H_{k-1}\left(\frac{c_1 - a}{\sqrt{c_2}}\right) \right) + \sqrt{c_2} H_{k-2}\left(\frac{c_1 - a}{\sqrt{c_2}}\right) \Phi\left(\frac{c_1 - a}{\sqrt{c_2}}\right). \end{aligned}$$

En appliquant cette méthode de Gram Charlier avec la variable $Y_T = B(T, T_N) + K \sum_{i=1}^N \delta B(T, T_i)$, il est possible de calculer les prix de swaptions facilement. Concernant l'évaluation des swaptions, la méthode de Gram Charlier fournit les meilleurs résultats à l'ordre 3 alors que pour le développement de type Edgeworth, il est préférable d'aller jusqu'à l'ordre 7. Concernant les méthodes d'expansion pour le calcul de swaptions, elles fournissent empiriquement de bonnes approximations pour les swaptions dont le tenor est inférieur à 10 ans.

Évaluation des swaptions par FFT

L'évaluation des swaptions peut se faire de manière classique sous la probabilité forward neutre en considérant le taux de swap comme un panier de zéro-coupon.

$$\begin{aligned}
 Sw(T, T_0, T_N, \delta, K) &= B(0, T) \mathbb{E}^{\mathbb{Q}_T} [(e^{Tr[A(T, T_0)V_T] + H^\top(T, T_0)X_T + C(T, T_0)} \\
 &\quad - e^{Tr[\bar{A}(T)V_T] + \bar{H}^\top(T)X_T + \bar{C}(T)})] \\
 &= B(0, T) \mathbb{E}^{\mathbb{Q}_T} [e^{Tr[A(T, T_0)V_T] + H^\top(T, T_0)X_T + C(T, T_0)} \\
 &\quad \cdot 1_{\{Tr[A(T, T_0)V_T] + H^\top(T, T_0)X_T + C(T, T_0) \geq Tr[\bar{A}(T)V_T] + \bar{H}^\top(T)X_T + \bar{C}(T)\}}] \\
 &\quad - B(0, T) \mathbb{E}^{\mathbb{Q}_T} [e^{Tr[\bar{A}(T)V_T] + \bar{H}^\top(T)X_T + \bar{C}(T)} \\
 &\quad \cdot 1_{\{Tr[A(T, T_0)V_T] + H^\top(T, T_0)X_T + C(T, T_0) \geq Tr[\bar{A}(T)V_T] + \bar{H}^\top(T)X_T + \bar{C}(T)\}}].
 \end{aligned}$$

Il faut donc calculer une fonction s'écrivant :

$$\Delta(T, A, B, C, D, E, F) = \mathbb{E}^{\mathbb{Q}_T} [\exp \{Tr[AV_T] + BX_T + C\} 1_{\{Tr[DV_T] + EX_T + F \geq 0\}}].$$

Cette fonction se calcule par transformée de Fourier en utilisant la transformée de Laplace des facteurs Φ .

3.2.2 Évaluation des caplets

Un caplet de maturité $T + \nu$ est une option sur le forward du taux Libor $L_t(T, T + \nu)$ à la date t , qui est fixé à la date T et qui paie en date $T + \nu$. Le caplet donne le droit d'entrée dans cette option au prix d'exercice K et peut être vue comme une swaption sur un swap à un flux de tenor δ . Son expression est donnée par :

$$\begin{aligned}
 C(t, T, K, \nu) &= B(0, T + \nu) \mathbb{E}_0^{\mathbb{Q}_{T+\nu}} [(L_t(T, T + \nu) - K)^+], \\
 L_t(T, T + \nu) &= \frac{1}{\nu} \left(\frac{B(t, T)}{B(t, T + \nu)} - 1 \right).
 \end{aligned}$$

Comme le cadre est affine, la méthode d'évaluation par FFT est assez classique. La transformée de Laplace $\Psi(\gamma, t, T, \nu)$ du log du prix du sous-jacent est donnée par

$$\Psi(\gamma, t, T, \nu) = \mathbb{E}_0^{\mathbb{Q}_{T+\nu}} [\exp \{ \gamma \log(1 + \nu L_t(T, T + \nu)) \}].$$

On peut exprimer le prix des caplets en utilisant l'expression de la transformée de Laplace. En effet, on peut établir la relation suivante :

$$\begin{aligned}
 \Psi(\gamma, t, T, \nu) &= \mathbb{E}_0^{\mathbb{Q}_{T+\nu}} [\exp \{ \gamma \log(1 + \nu L_t(T, T + \nu)) \}] = \mathbb{E}_0^{\mathbb{Q}_{T+\nu}} \left[\exp \left\{ \gamma \log \left(\frac{B(t, T)}{B(t, T + \nu)} \right) \right\} \right] \\
 &= \frac{\exp \{ \gamma (C(t, T) - C(t, T + \nu)) \}}{\mathbb{E}_0^{\mathbb{Q}_{T+\nu}} [\exp \{ \gamma Tr[(A(t, T) - A(t, T + \nu))V_t] + (H^\top(t, T) - H^\top(t, T + \nu))X_t \}]} \\
 &= \frac{\exp \{ \gamma (C(t, T) - C(t, T + \nu)) \}}{\Phi(0, t, T + \nu, \gamma(A(t, T) - A(t, T + \nu)), \gamma(H(t, T) - H(t, T + \nu)), V_0, X_0)}. \tag{3.2}
 \end{aligned}$$

Par conséquent, la transformée de Laplace du sous-jacent Ψ s'exprime en fonction de la transformée de Laplace des facteurs Φ . Ainsi, le prix des caplets peut se déduire aisément par FFT. La méthode d'évaluation par FFT est classique pour les modèles affines et elle a déjà été utilisée dans le chapitre 2 de la thèse.

Une autre méthode d'évaluation est de considérer que le caplet est une swaption sur un swap avec une seule échéance de paiement. La méthode des moments peut donc être utilisée à la fois pour l'évaluation des caplets et l'évaluation des swaptions. De plus, cette méthode fournit des résultats assez précis [4].

3.3 Étude d'un modèle de Wishart équivalent à un modèle de Heston

Cette partie est consacrée à l'évaluation des caplets et des swaptions dans le modèle de Wishart. Tout d'abord, le modèle de Wishart se révèle être une extension multidimensionnelle du modèle de Heston puisqu'il est possible de spécifier le modèle de sorte à retrouver le modèle standard de Heston. Le modèle de Heston sert de référence car on connaît certains comportements classiques concernant les courbes de prix de zéro-coupon, les smiles de caplets et les surfaces de volatilité de swaptions. En fait, il existe une restriction un peu spéciale liée au processus de Wishart lui-même : lorsqu'on utilise le processus de Wishart pour lequel le terme de la volatilité à l'infini est sous la forme $\beta Q^\top Q$, la volatilité à l'infini est fortement liée à la matrice Q avec pour seul degré de liberté le paramètre β réel. Il est en fait possible de passer outre ce cadre en considérant un processus de Wishart plus général sous la forme :

$$dV_t = M(V_t - V_\infty) + (V_t - V_\infty)M^\top + \sqrt{V_t}dW_tQ + Q^\top dW_t^\top \sqrt{V_t}.$$

mais il est plus difficile d'interpréter le processus de Wishart comme une extension de la somme de carrés d'Ornstein Uhlenbeck et d'interpréter le paramètre β comme degré d'un khi-deux. Dans cette étude, la dynamique du processus de Wishart est supposée de la forme :

$$dV_t = \beta Q^\top Q + MV_t + V_tM^\top + \sqrt{V_t}dW_tQ + Q^\top dW_t^\top \sqrt{V_t}.$$

3.3.1 Présentation du modèle

Nous considérons un modèle de Heston en dimension 1 sur le taux court dont la dynamique s'écrit sous la forme suivante :

$$\begin{aligned} dr_t &= \lambda(l - r_t)dt + \sqrt{V_t}(\rho dW_t + \sqrt{1 - \rho^2}dB_t) \\ dV_t &= \kappa(\theta - V_t)dt + \epsilon\sqrt{V_t}dW_t \\ d\langle W, B \rangle &= 0. \end{aligned} \tag{3.3}$$

Le modèle de Wishart permettant de reproduire le modèle de Heston dans le cas où $\epsilon \neq 0$ et $\rho \neq 1$ est spécifié comme suit :

$$\begin{aligned} r_t &= l - \frac{\rho\kappa\theta}{\lambda\epsilon} + \frac{\rho}{\epsilon}V_t + \sqrt{1-\rho^2}X_t \\ dV_t &= \kappa(\theta - V_t)dt + \epsilon\sqrt{V_t}dW_t \\ dX_t &= -\lambda X_t dt + \frac{(\kappa-\lambda)\rho}{\epsilon\sqrt{1-\rho^2}}V_t dt + \sqrt{V_t}dB_t. \end{aligned} \quad (3.4)$$

On se place dans le cadre où la courbe des taux est croissante. Partant d'un taux court $r_0 = 2\%$, on considère une courbe des taux croissante de sorte que $l = 4\%$ avec une volatilité initiale sur le taux court de l'ordre de 1%. L'ensemble suivant de paramètre prend en compte cette configuration en ajoutant les notions de retour à la moyenne et de volatilité pour le taux court et sa volatilité :

$$\lambda = 0.1, \quad l = 4\%, \quad r_0 = 2\%, \quad \kappa = 8\%, \quad \theta = 10^{-4}, \quad \epsilon = 10^{-4}, \quad \rho = -0.2. \quad (3.5)$$

Spécifiant le modèle (3.4) de Wishart équivalent, on teste le comportement du modèle. Tout d'abord, on analyse la courbe de prix des zéro-coupon ainsi que la courbe des taux générée par le modèle.



FIGURE 3.1 – Courbe de prix de zéro-coupon générée par le modèle

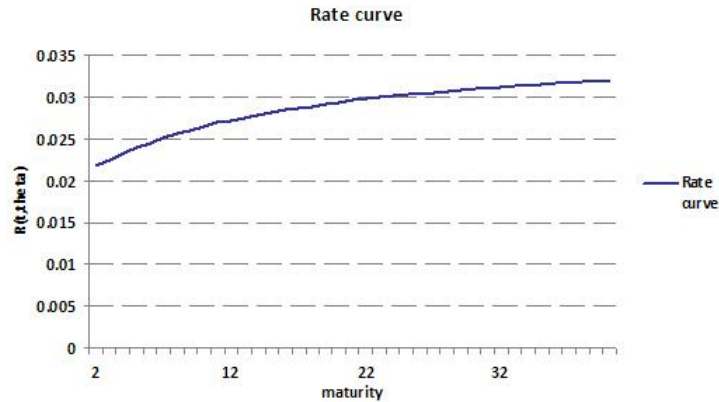


FIGURE 3.2 – Courbe de taux générée par le modèle

3.3.2 Smile de caplets et surface de volatilité à la monnaie (ATM)

Les concepts de smile de caplets et de surface de volatilité à la monnaie (ATM) qui sont les informations principales de marché disponibles. Nous analysons les smiles de caplets et les surfaces de volatilité de swaptions que le modèle peut générer. Ces instruments permettent de calibrer les modèles utilisés dans la couverture de produits exotiques.

Volatilité implicite

Les prix des options sur le marché dépendent de certaines caractéristiques comme le prix d'exercice (strike), la maturité et éventuellement le tenor sur le marché des taux d'intérêt. La volatilité implicite lognormale (ou Black) correspond à la volatilité du sous-jacent qu'il faut considérer pour retrouver le prix de l'option en supposant que ce dernier suit une dynamique lognormale. La volatilité implicite de normale (ou Bachelier) correspond à la volatilité du sous-jacent dans le cas d'une dynamique normale. Par conséquent, dans le cas d'une swaption, la volatilité implicite Bachelier suppose une dynamique normale du taux de swap et la volatilité implicite Black suppose une dynamique lognormale du taux de swap. La volatilité implicite normale est assez standard dans le monde des taux d'intérêt dans la mesure où son ordre de grandeur est plus intuitif. En effet, il est plus aisé d'interpréter la valeur de la volatilité d'un taux lorsque ce dernier est suivi d'une dynamique normale. Pour plus de lisibilité, les volatilités implicites sont souvent exprimées et représentées en points de base ($1BP = 10^{-4} = 0.01\%$).

La volatilité ATM correspond au niveau de volatilité implicite lorsque le prix de l'option est égal au taux forward sous-jacent. Lorsque le prix d'exercice de l'option est différent du taux forward sous-jacent, la volatilité de l'option est généralement différente de celle constatée ATM. Ce constat est appelé "effet smile" dont l'étude est un enjeu majeur de marché. En

effet, les modèles cherchent à capter au mieux ce phénomène afin de reproduire au mieux les prix d'options vanilles pour l'ensemble des dates et des prix d'exercice.

Smile de caplets

Les méthodes numériques permettent d'évaluer les prix de caplets. Un caplet peut être vu comme un cas particulier d'une swaption et la méthode des moments permet de calculer les prix de caplets de maturité T et de prix d'exercice K associé au Libor $L(T, T + \delta)$ fixé en T pour la maturité $T + \delta$. On définit un caplet $T_1 * \delta$ au prix d'exercice K comme l'option d'achat de rentrer à la maturité T_1 sur un taux Libor $L(T_1, T_1 + \delta)$ au prix d'exercice K .

À partir des prix de caplets, il est possible de considérer la volatilité implicite du sous-jacent qui permet de récupérer le prix obtenu. La structure de la volatilité implicite en maturité et en strike fournit les smiles de caplets. Pour l'ensemble de paramètre donné en 3.5, on représente un exemple de smiles en fonction du prix d'exercice pour différentes maturités.

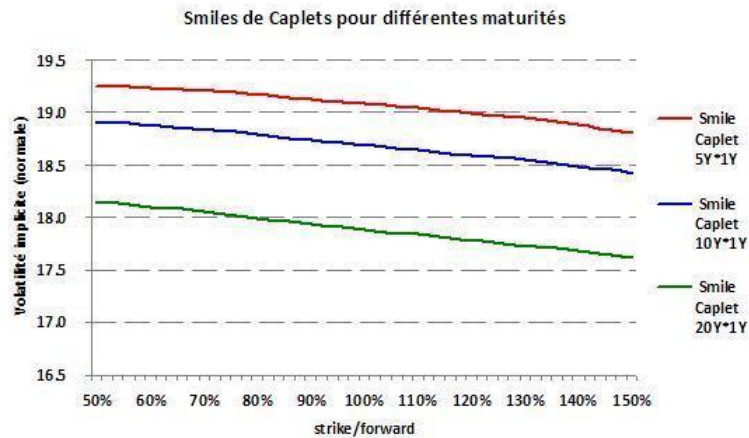


FIGURE 3.3 – Smile de caplets sur le Libor 1 an pour différentes maturités (en BP)

Le niveau du smile de caplets est généralement une fonction décroissante de la maturité de l'option. Les formes de smiles sont assez restreintes en dimension 1 mais il est tout de même possible d'atteindre des configurations assez variées.

Prix et volatilité implicite des swaptions

On présente une étude rapide des prix de swaptions et de leur surface de volatilité implicite ATM. On analyse tout d'abord l'influence de la maturité T et du tenor $T_N - T$ sur les prix de swaptions. Il est standard et pertinent de considérer le prix "renormalisé" de la swaption

par le "level" de la swaption qui est donné par :

$$Level_t(T, T_N) = \sum_{i=1}^N \delta B_t(T, T_i).$$

En renormalisant le prix de swaption, il est plus commode d'interpréter l'impact des paramètres de l'option. En effet, en considérant $Level_0(T, T_N)$ comme numéraire ce qui revient à considérer le changement de probabilité suivant :

$$\frac{d\mathbb{Q}_{LV L}}{d\mathbb{Q}_T}\bigg|_t = \frac{B(t, T)}{Level_t(T, T_N)} \sum_{i=1}^N e^{-\int_T^{T_i} r_s ds},$$

il est possible d'exprimer le prix renormalisé de la swaption sous la forme :

$$\frac{Sw(T, T_0, T_N, \delta, K)}{Level_0(T, T_N)} = \mathbb{E}^{\mathbb{Q}_{LV L}} [(S_T(T_0, T_N, \delta) - K)^+]$$

Le prix renormalisé est fonction du taux de swap forward qui est martingale sous la probabilité $\mathbb{Q}_{LV L}$, et on pourra plus facilement interpréter les applications numériques. On observe que le prix renormalisé d'une swaption est une fonction croissante de la maturité quelque soit le tenor. On peut d'ailleurs représenter l'évolution du prix de swaptions en fonction de la maturité pour plusieurs tenors :

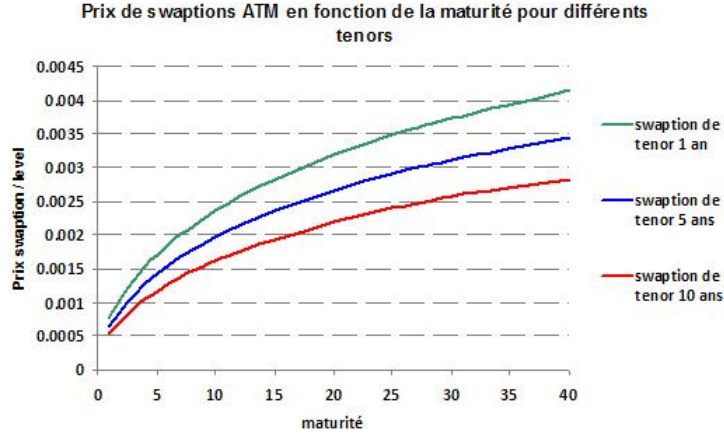


FIGURE 3.4 – Prix de swaptions en fonction de la maturité pour plusieurs tenor

Ce résultat n'est pas forcément évident à interpréter puisque le taux de swap augmente aussi avec la maturité. Cependant, on sait que le prix d'une option permettant de rentrer dans un swap fixe est une fonction croissante de la maturité étant donné le risque de volatilité

augmentant dans le temps. Le modèle fournit des prix de ces options qui vérifient cette propriété (voir Figure 3.5).

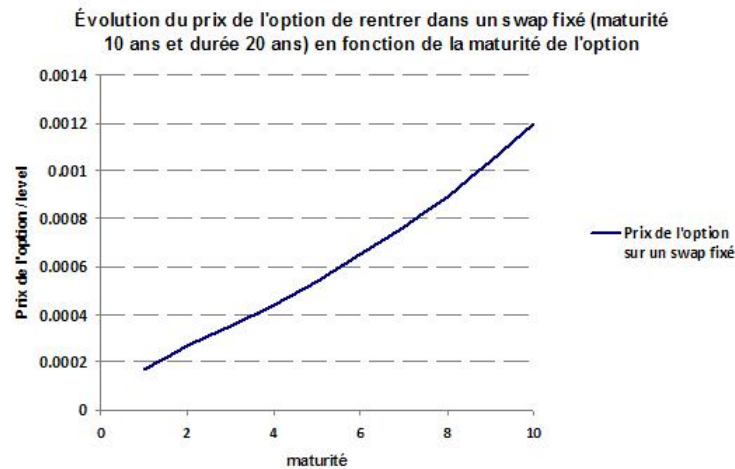


FIGURE 3.5 – Prix d'une option sur un swap fixe en fonction de la maturité

On regarde maintenant l'influence du tenor sur le prix des swaptions. Pour plusieurs maturités, on a calculé le prix de swaptions renormalisé par le level en fonction du tenor pour plusieurs maturités. On constate le comportement classique d'un prix de swaption décroissant du tenor.

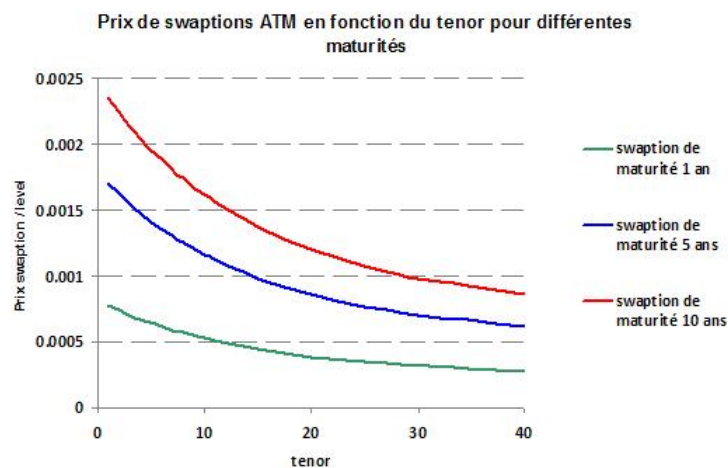


FIGURE 3.6 – Prix de swaptions en fonction du tenor pour plusieurs maturités

Enfin, on génère la surface de volatilité ATM associée aux prix de swaptions pour différentes maturités et différents tenors. Cette configuration est assez standard sur les marchés avec

une forte décroissance en tenor et une décroissance moindre en maturité.

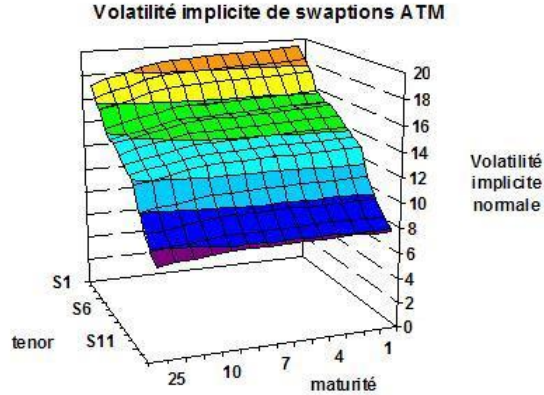


FIGURE 3.7 – Surface de volatilité implicite de swaption ATM en fonction de la maturité et le tenor (en BP)

On propose une étude très qualitative des impacts des paramètres sur les surfaces de volatilité et les smiles de caplets. Cela sera assez utile pour un trader qui souhaite trouver un ensemble de paramètres permettant d'obtenir des prix de marché vraisemblables. Cette analyse de l'impact des paramètres sur les surfaces de volatilité de swaptions et sur le smile de caplets est bien évidemment assez pauvre et il faudra effectuer des tests plus poussés avant d'utiliser le modèle.

3.3.3 Influence des paramètres du modèle

Nous nous intéressons aux effets d'un mouvement des paramètres sur les smiles de caplets et sur les surfaces de volatilité. Quelques phénomènes sont intéressants à relever permettant de comprendre l'influence des paramètres. Pour cela, nous considérons une courbe initiale des taux croissante ainsi qu'un processus de variance peu explosif en supposant $V_\infty \sim V_0$. D'autres configurations pourraient évidemment être étudiées de manière similaire.

En dehors de la corrélation ρ qui contrôle la pente du smile de caplets en fonction du prix d'exercice (skew), on teste l'influence d'un mouvement des paramètres sur la surface de volatilité des swaption à la monnaie. On fixe une valeur initiale raisonnable des paramètres, on considère un choc additif de 10% d'un des paramètres. On regarde l'écart qu'il existe entre la surface de volatilité déformée avec la surface de volatilité initiale. On parlera d'une *influence positive* si l'augmentation de la valeur du paramètre entraîne une montée du niveau de volatilité implicite et d'une *influence négative* si elle engendre une baisse du niveau de volatilité implicite.

Influence du paramètre ρ

On peut par exemple regarder l'influence du paramètre ρ sur les prix de caplets afin de voir si le modèle a le comportement attendu. Le smile de caplets est fortement influencé par le paramètre de corrélation ρ et en particulier que le signe de la pente est contrôlé par le signe de ρ .

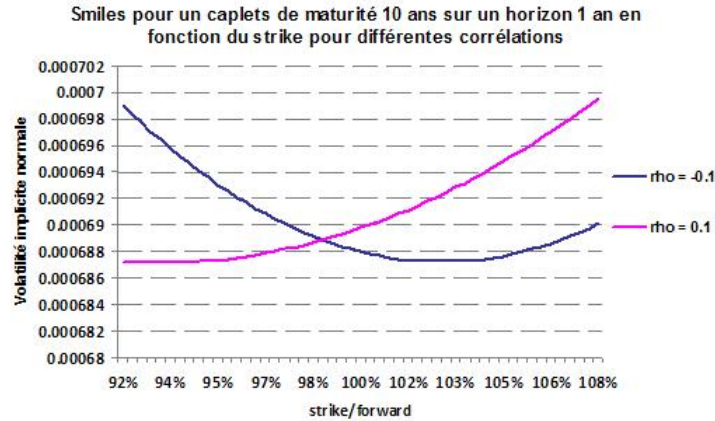


FIGURE 3.8 – Influence de ρ sur le smile de caplets ($T = 10Y$, $U = 11Y$)

La corrélation ρ influence aussi le niveau global de la surface de volatilité implicite ATM mais n'a pas d'impact localisé en fonction de la maturité ou du tenor de la swaption comme les autres paramètres du modèle.

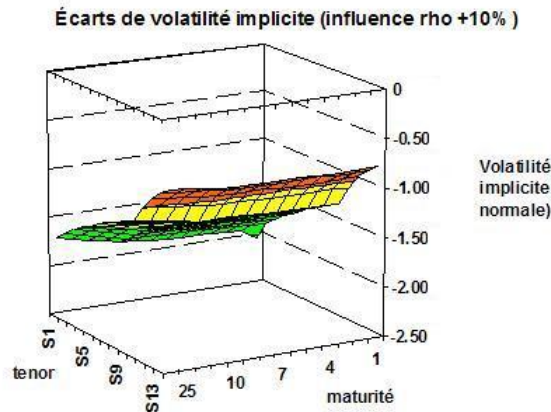


FIGURE 3.9 – Influence de ρ sur la surface de volatilité ATM (en BP)

Comme le prévoyait le modèle de Heston, le paramètre ρ pilote le signe du skew de smile de caplets et a une influence globale sur le niveau de la surface de volatilité des swaptions

ATM.

Remarque 3.1. *La corrélation ρ est souvent supposée strictement négative puisqu'on observe que lorsque les taux augmentent, leur volatilité diminue et inversement. Ainsi, on observe généralement un smile de caplets décroissant en strike. Plus la corrélation est petite en valeur absolue, plus le smile de caplets est convexe.*

Influence du paramètre λ

Le paramètre λ est le paramètre de vitesse de convergence de la moyenne du taux court vers sa valeur à l'infini. Plus le paramètre λ est grand, plus l'effet de retour à la moyenne est intense ce qui fait que la moyenne du taux court converge rapidement vers sa valeur à l'infini l . On observe la déformation de la surface de volatilité résultant d'une augmentation du paramètre λ .

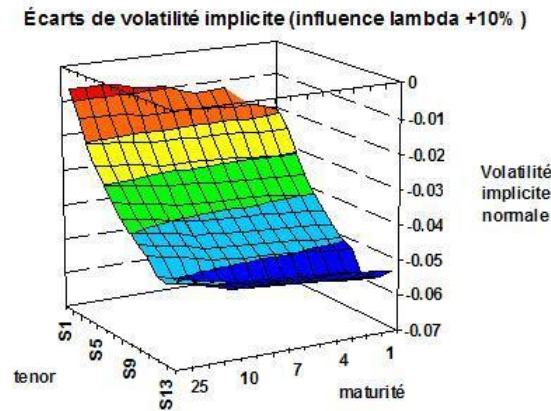


FIGURE 3.10 – Influence de λ sur la surface de volatilité ATM (en BP)

Ce paramètre a une "influence négative" assez faible qui augmente à mesure que les tenors augmentent. L'influence de ce paramètre dépend très peu de la maturité de la swaption.

Influence du paramètre θ

Le paramètre θ correspond au niveau moyen de la volatilité à l'infini. Plus le paramètre θ augmente, plus le niveau long terme de la volatilité va être important. Lorsque le paramètre θ est de l'ordre de V_0 , la volatilité est contrainte à rester assez proche de son niveau initial.

Le paramètre θ a une "influence positive" importante sur la surface de volatilité localisée sur les grandes maturités. Ce comportement est assez intuitif puisque ce paramètre joue sur le niveau long terme de la volatilité.

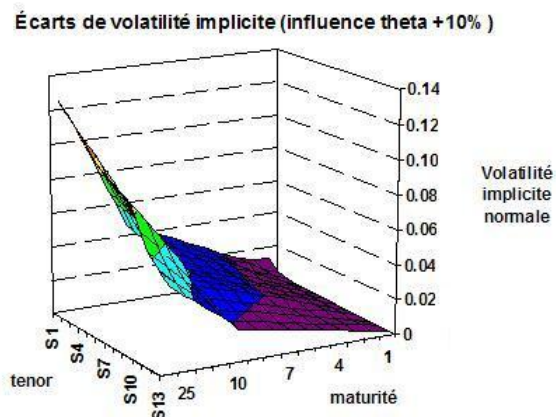


FIGURE 3.11 – Influence de θ sur la surface de volatilité ATM (en BP)

Influence du paramètre κ

Le paramètre κ est le paramètre de vitesse de convergence de la moyenne de la volatilité vers sa valeur à l'infini. Plus le paramètre κ est grand, plus on se rapproche du cas où la moyenne de la volatilité converge rapidement vers θ . Les rôles de κ et de θ sont donc étroitement liés. On observe la déformation de la surface de volatilité résultant d'une augmentation du paramètre κ .

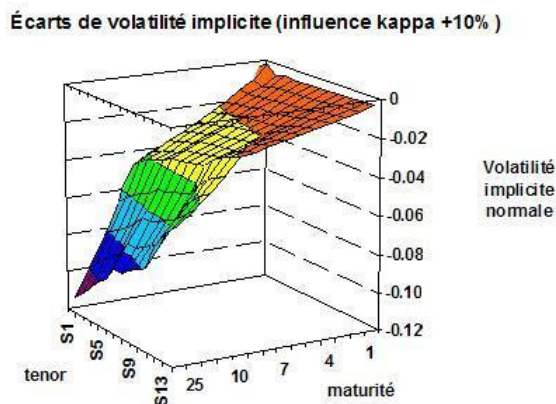


FIGURE 3.12 – Influence de κ sur la surface de volatilité ATM (en BP)

Le paramètre κ a une "influence négative" qui s'intensifie à mesure que la maturité augmente. Alors que le paramètre λ de vitesse de retour à la moyenne du taux a un impact important pour les grands tenors, le paramètre κ a lui un impact sur les options de grandes maturités. Ainsi, on observe des rôles assez dissymétriques des paramètres pilotant la dynamique du taux par rapport à ceux contrôlant la dynamique de la volatilité.

Influence du paramètre ϵ

Le paramètre ϵ est le paramètre de volatilité de la volatilité du taux court. Un paramètre ϵ autorise cette volatilité à varier plus ou moins brutalement autour de la valeur θ .

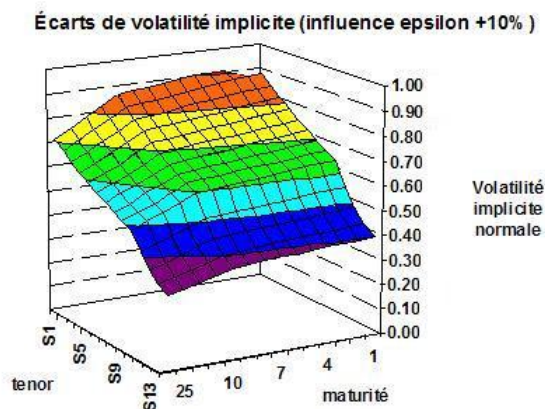


FIGURE 3.13 – Influence de ϵ sur la surface de volatilité ATM (en BP)

Le paramètre ϵ de volatilité de la volatilité a une "influence positive". On remarque que l'influence est importante pour les petites maturités et/ou les petits tenors. Cette influence est assez spéciale dans la mesure où c'est le seul paramètre qui joue à la fois sur les petits tenors et les petites maturités. Ainsi, pour des temps petits, ce paramètre a une influence importante puisque les dynamiques des processus n'ont pas encore atteint leur régime long terme.

Bilan sur l'influence des paramètres

Cette étude révèle une influence variée des paramètres sur la matrice de volatilité : l'augmentation de certains paramètres peut soit faire baisser le niveau de volatilité soit l'augmenter. Nous remarquons que les paramètres de vitesse de retour à la moyenne (λ et κ) ont un impact négatif sur le niveau de la surface de volatilité ATM alors que les autres ont un impact positif.

De plus, nous constatons que l'impact n'est pas uniforme et que chaque paramètre a une influence localisée sur la surface de volatilité. Les paramètres ont des influences variées et nous pouvons essayer d'adapter les paramètres de manière à ajuster la surface de volatilité sur l'ensemble des maturités et tenors. D'ailleurs, nous observons des comportements assez différents des paramètres de la dynamique du taux par rapport à ceux de la dynamique de la volatilité.

3.3.4 Étude d'un modèle Wishart en dimension 2

Pour information, nous décrivons le modèle équivalent à un modèle de Heston à deux facteurs. La modélisation classique du modèle de Heston s'écrit :

$$\begin{aligned} dr_t &= \lambda(l - r_t)dt + \sqrt{V_t^1}(\rho_1 dW_t + \sqrt{1 - \rho_1^2}dB_t) + \sqrt{V_t^2}(\rho_2 dW_t + \sqrt{1 - \rho_2^2}dB_t) \\ dV_t^1 &= \kappa_1(\theta_1 - V_t)dt + \epsilon_1 \sqrt{V_t} dW_t \\ dV_t^2 &= \kappa_2(\theta_2 - V_t)dt + \epsilon_2 \sqrt{V_t} dW_t \\ d < W, B > &= 0. \end{aligned}$$

Le modèle de Wishart peut dégénérer en un modèle de Heston de dimension 2. La spécification du modèle peut s'écrire de la manière suivante :

$$\begin{cases} dX_t = P(\mathbb{X} - X_t)dt + \begin{pmatrix} Tr(\Pi_1 V_t) \\ \vdots \\ Tr(\Pi_p V_t) \end{pmatrix} dt + R\sqrt{diag(\alpha_i + Tr(\Sigma_i V_t) : 1 \leq i \leq h)}d\vec{Z}_t \\ dV_t = (\beta Q^\top Q + MV_t + V_t M^\top)dt + \sqrt{V_t}dW_t Q + Q^\top dW_t^\top \sqrt{V_t} \\ r_t = \alpha + Tr[\Sigma V_t] + \delta^\top X_t, \end{cases}$$

avec

$$\begin{aligned} P &= \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}, \Pi_1 = \begin{pmatrix} \frac{\rho_1}{\epsilon_1}(\lambda - \kappa_1) & 0 \\ 0 & 0 \end{pmatrix}, \Pi_2 = \begin{pmatrix} 0 & 0 \\ 0 & \frac{\rho_2}{\epsilon_2}(\lambda - \kappa_2) \end{pmatrix}, \mathbb{X} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ \Sigma_1 &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \Sigma_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \alpha_1 = \alpha_2 = 0, \\ Q &= \begin{pmatrix} \frac{\epsilon_1}{2} & 0 \\ 0 & \frac{\epsilon_2}{2} \end{pmatrix}, M = \begin{pmatrix} -\kappa_1 & 0 \\ 0 & -\kappa_2 \end{pmatrix}, \beta = \frac{4\kappa_1\epsilon_1}{\epsilon_1^2} = \frac{4\kappa_2\epsilon_2}{\epsilon_2^2}, \\ \alpha &= l - \frac{\rho_1\kappa_1\theta_1}{\lambda\epsilon_1} - \frac{\rho_2\kappa_2\theta_2}{\lambda\epsilon_2}, \Sigma = \begin{pmatrix} \frac{\rho_1}{\epsilon_1} & 0 \\ 0 & \frac{\rho_2}{\epsilon_2} \end{pmatrix} \delta_1 = \sqrt{1 - \rho_1^2}, \delta_2 = \sqrt{1 - \rho_2^2}. \end{aligned}$$

Remarque 3.2. La contrainte sur le paramètre β du processus de Wishart se traduit par une contrainte sur les paramètres du modèle de Heston. Il sera préférable d'utiliser une forme générale pour la matrice $\Omega^\top \Omega$ de manière à conserver une flexibilité sur la volatilité à l'infini.

Ainsi, le modèle de Wishart peut généraliser les modèles de Heston multidimensionnel. Nous avons vu au chapitre 2 que le modèle de Wishart dans le monde des actions était plus flexible que le modèle de Heston avec un degré de liberté sur la corrélation stochastique. Cette flexibilité pourra être utile pour capter certains phénomènes comme par exemple les formes de smiles de caplets, les formes de surfaces de volatilité de swaptions et la dynamique du skew.

N'ayant pas encore mis en place une procédure de calibration de ce modèle affine sur les prix des options vanilles, nous avons testé le modèle avec un ensemble vraisemblable de paramètres. En vue d'une procédure de calibration efficace, il faudra trouver des formules d'approximation de smile de caplets et de surface de volatilité.

3.4 Propriétés intrinsèques du modèle de Wishart

Dans cette partie, nous comparons le modèle de Wishart à un modèle affine standard. La structure de volatilité du modèle de Wishart étant quadratique, nous introduisons une projection du modèle sur l'ensemble des modèles affines.

De plus, nous étudions une spécification du modèle permettant de capter la corrélation du taux court avec sa volatilité qui contrôle la dynamique du skew des options sur taux court. Avec une telle spécification, il est possible le risque du skew d'options sur taux court qui est considéré comme un risque exotique.

3.4.1 Comparaison modèle Wishart/ modèle affine (Daï-Singleton)

Le modèle de Wishart peut être spécifié comme un modèle affine de type Daï & Singleton du fait de la présence du processus de Wishart qui confère aussi un aspect quadratique au modèle. Les modèles affines de type Daï-Singleton (DS) s'écrivent sous la forme suivante :

$$\left\{ \begin{array}{l} dY_t = (u - v^\top Y_t)dt + \Sigma \sqrt{S_t} dW_t \\ (S_t)^{ii} = \alpha_i + \sum_{k=1}^n \beta_{ki} Y_t^k, \quad (S_t)^{ij} = 0 \text{ pour } i \neq j \\ r_t = \delta_0 + \delta_1^\top Y_t. \end{array} \right.$$

Pour faire le lien entre le modèle de Wishart et les modèles de DS, il faut décrire l'ensemble des facteurs du modèle de Wishart : ce dernier contient $n^2 + h$ facteurs et il faut les ordonner afin de les numéroter.

L'ensemble des facteurs peut être décrit sous la forme d'un vecteur Y_t de taille $n^2 + h$ donné par :

$$Y_t = \begin{pmatrix} X_t \\ \tilde{V}_t \end{pmatrix}, \text{ avec } \tilde{V}_t = (V_t^{11}, \dots, \Pi_i^{1n}, \dots, \Pi_i^{n1}, \dots, \Pi_i^{nn}).$$

Correspondance des structures Wishart et affines

Un petit calcul rapide fournit un lien explicite en ce qui concerne le drift. La correspondance est donnée par :

$$u = \begin{matrix} h \downarrow \\ n^2 \uparrow \end{matrix} \left(\begin{array}{c} P\mathbb{N} \\ \beta \tilde{Q}^\top Q \end{array} \right), \quad C = \begin{pmatrix} M^{11}I_n & \dots & M^{1n}I_n \\ \vdots & & \vdots \\ M^{n1}I_n & \dots & M^{nn}I_n \end{pmatrix}, \quad v^\top = \begin{matrix} h \downarrow \\ n^2 \uparrow \end{matrix} \left(\begin{array}{cc} \xleftrightarrow{h} & \xleftrightarrow{n^2} \\ -P & \Pi \\ 0 & C \end{array} \right),$$

$$\Pi = \begin{pmatrix} \Pi_1 \\ \vdots \\ \Pi_h \end{pmatrix} \quad \text{où} \quad \Pi_i = (\Pi_i^{11}, \dots, \Pi_i^{n1}, \dots, \Pi_i^{1n}, \dots, \Pi_i^{nn}).$$

En ce qui concerne la volatilité, il existe des différences entre la structure affine et la structure Wishart.

En effet, on peut écrire :

$$\Sigma^{Wishart} = \begin{matrix} h \downarrow \\ n^2 \uparrow \end{matrix} \left(\begin{array}{cc} \xleftrightarrow{h} & \xleftrightarrow{n^2} \\ R\sqrt{D_0} & 0 \\ 0 & A \end{array} \right), \quad \Sigma^{DS} = \Sigma\sqrt{S},$$

avec A , qui correspond à la projection de $\sqrt{V_t}W_tQ + Q^\top W_t^\top \sqrt{V_t}$ sur l'ensemble des facteurs. Cette matrice de taille $n^2 \times n^2$ a une forme compliquée (de taille $n^2 \times n^2$) dont on décrira juste quelques lignes pour information :

$$A = \begin{matrix} \begin{matrix} i=1 \\ j=1 \end{matrix} \rightarrow \\ \begin{matrix} i=1 \\ j=2 \end{matrix} \rightarrow \\ \\ \begin{matrix} i=2 \\ j=1 \end{matrix} \rightarrow \end{matrix} \left(\begin{array}{cccc|cccc|ccc} 2Q^{11}\sqrt{V_t^{11}} & \dots & 2Q^{11}\sqrt{V_t^{1n}} & & 2Q^{21}\sqrt{V_t^{11}} & \dots & 2Q^{21}\sqrt{V_t^{1n}} & & \dots \\ Q^{12}\sqrt{V_t^{11}} + Q^{11}\sqrt{V_t^{12}} & \dots & Q^{12}\sqrt{V_t^{1n}} + Q^{11}\sqrt{V_t^{n2}} & & Q^{22}\sqrt{V_t^{11}} + Q^{21}\sqrt{V_t^{12}} & \dots & Q^{22}\sqrt{V_t^{1n}} + Q^{21}\sqrt{V_t^{n2}} & & \dots \\ \vdots & & & & \vdots & & & & \dots \\ Q^{11}\sqrt{V_t^{21}} + Q^{12}\sqrt{V_t^{11}} & \dots & Q^{11}\sqrt{V_t^{2n}} + Q^{12}\sqrt{V_t^{n1}} & & Q^{21}\sqrt{V_t^{21}} + Q^{22}\sqrt{V_t^{11}} & \dots & Q^{21}\sqrt{V_t^{2n}} + Q^{22}\sqrt{V_t^{1n}} & & \dots \\ \vdots & & & & \vdots & & & & \dots \end{array} \right)$$

On peut écrire les matrices du modèle Dai-Singleton sous la forme :

$$\Sigma^{DS} = \begin{matrix} h \downarrow \\ n^2 \uparrow \end{matrix} \left(\begin{array}{cc} \xleftrightarrow{h} & \xleftrightarrow{n^2} \\ \Sigma_1^{DS} & 0 \\ 0 & \Sigma_2^{DS} \end{array} \right), \quad S_t = \begin{matrix} h \downarrow \\ n^2 \uparrow \end{matrix} \left(\begin{array}{cc} \xleftrightarrow{h} & \xleftrightarrow{n^2} \\ (S_1)_t & 0 \\ 0 & (S_2)_t \end{array} \right).$$

Pour la partie affine, la résolution ne pose aucun problème, elle est donnée par $R\sqrt{D_0} = \Sigma_1^{DS}\sqrt{(S_1)_t}$ ce qui s'écrit aussi $\Sigma_1^{DS} = R$ et $(S_1)_t = D_0$.

En revanche, en ce qui concerne la partie Wishart, il n'est pas possible d'écrire $A =$

$\Sigma_2^{DS} \sqrt{(S_2)_t}$. Nous proposons par conséquent une sorte de projection de notre composante Wishart sur l'espace des modèles affines, c'est à dire de trouver les Σ_2^{DS} et $(S_2)_t$ optimaux qui approchent au mieux la matrice A .

En effet, la projection est caractérisée par la relation suivante :

$$\begin{aligned} \|r_t^{Wishart} - r_t^{DS}\|_2^2 &= \mathbb{E} \left[\left(\int_0^T (\Sigma_2^{DS} \sqrt{S_t} - \Sigma^{Wishart}) dW_t \right)^\top \left(\int_0^T (\Sigma_2^{DS} \sqrt{S_t} - \Sigma^{Wishart}) dW_t \right) \right] \\ &= \mathbb{E} \left(\int_0^T \|\Sigma_2^{DS} \sqrt{(S_2)_t} - A\|_2^2 dt \right). \end{aligned}$$

Proposition de projection du modèle de Wishart dans un cadre affine

Il n'apparaît pas évident de trouver les matrices $(S_2)_t$ (devant être diagonale) et Σ_2^{DS} optimales. En revanche, on peut remarquer que si V_t restait diagonale dans le temps (ce qui n'est pas possible dans une structure Wishart même si M et Q sont diagonales), les matrices Σ_2^{DS} et $(S_2)_t$ optimales seraient facilement identifiables. Un choix naturel est alors de considérer deux matrices de sorte à annuler tous les termes de la matrice A faisant apparaître les composantes diagonales de V_t correspondant à l'ensemble des V_t^{ii} . Par conséquent, $A - \Sigma_2^{DS} \sqrt{(S_2)_t}$ ne dépendrait que des V_t^{ij} avec $i \neq j$.

Il est facile de voir que le bon choix se ramène aux matrices suivantes :

$$\begin{cases} (\Sigma_2^{DS})^{in+j,ln+k} = 1_{\{k=i\}} Q^{lj} + 1_{\{k=j\}} Q^{li} \\ (S_2)_t^{in+j,in+j} = V_t^{jj} \quad , \quad (S_2)_t^{in+j,kn+l} = 0 \quad \text{si } k \neq i \text{ ou } l \neq j \end{cases}$$

En effet, ces deux matrices fournissent une projection du Wishart sur l'espace affine associé permettant de ne considérer uniquement les composantes diagonales de V_t . Cette projection a d'autant plus de sens dans le cas où les composantes non diagonales de V_t^{ij} sont petites.

3.4.2 Spécification d'un modèle qui contrôle la dynamique du skew

La dynamique du skew est souvent considérée comme un risque exotique et le skew pourrait devenir un risque vanille qu'il peut être possible de couvrir. En effet, ce modèle permet de contrôler le skew par les facteurs du modèle. La présence du facteur X_t sera justifiée dans la dynamique du taux court.

Étude du modèle pour des options sur taux court

La volatilité du taux court s'exprime en fonction des composantes de la matrice de Wishart. On exhibe un modèle (en dimension $n = 2$) dans lequel le skew peut être contrôlé par une composante de la matrice de Wishart V_t n'apparaissant pas dans l'expression de la volatilité. On traite dans un premier temps le cas d'une option sur taux court r_t . En

considérant la formule de Bachelier ainsi qu'une volatilité implicite normale, la volatilité du taux court peut être vue comme l'autocorrélation du taux court $d < r >_t$ et le skew peut être représenté (à un facteur de renormalisation près) comme la corrélation entre le taux court et sa variance $d < r, Vol^2 >_t$. Un petit calcul matriciel préliminaire donne les formules utiles suivantes :

$$d < Tr(AdWB) >_t = Tr(BAA^\top B^\top)dt, \quad (3.6)$$

$$d < Tr(AdWB), Tr(CdWD) >_t = Tr(DCA^\top B^\top)dt. \quad (3.7)$$

On pose $a_k = \sum_{i=1}^p \delta^i R^{ik}$. En utilisant les formules (3.6) et (3.7), on peut calculer l'auto-corrélation du taux court ainsi que la corrélation entre le taux court et sa volatilité :

$$\begin{aligned} d < r >_t &= [2Tr(\Sigma Q^\top Q \Sigma V_t) + \sum_{k=1}^h (\alpha_k + Tr(\Sigma_k V_t)) a_k^2] dt, \\ d < r, V^{ij} >_t &= [(V_t \Sigma Q^\top Q)^{ij} + (V_t \Sigma Q^\top Q)^{ji}] dt. \end{aligned}$$

On se concentre sur la quantité $d < r, Vol^2 >_t$ où Vol_t représente la volatilité du taux court à l'instant t . Le cadre affine permet d'écrire Vol_t^2 comme combinaison linéaire de tous les facteurs :

$$Vol_t^2 = a_1(t) V_t^{11} + a_2(t) V_t^{22} + a_3(t) V_t^{12}.$$

Ainsi, on déduit l'expression du skew (à un coefficient multiplicateur près) :

$$Skew \sim d < r, Vol^2 >_t = [b_1(t) V_t^{11} + b_2(t) V_t^{22} + b_3(t) V_t^{12}] dt,$$

dont les coefficients sont donnés par les équations suivantes :

$$\begin{aligned} b_1(t) &= 2(\Sigma Q^\top Q)^{11} a_1(t) + 2(\Sigma Q^\top Q)^{12} a_3(t), \\ b_2(t) &= 2(\Sigma Q^\top Q)^{22} a_2(t) + 2(\Sigma Q^\top Q)^{21} a_3(t), \\ b_3(t) &= 2(\Sigma Q^\top Q)^{21} a_1(t) + 2(\Sigma Q^\top Q)^{12} a_2(t) + [(\Sigma Q^\top Q)^{22} + (\Sigma Q^\top Q)^{11}] a_3(t). \end{aligned}$$

Ainsi, nous remarquons que n'importe quelle composante de la volatilité fournit a priori une composante sur le skew en V_t^{12} . Cela nous amène à penser que s'il existe un modèle dans lequel le skew est porté par un seul facteur, ce facteur serait V_t^{12} .

En étudiant l'ensemble des cas possibles, on se rend compte que dans le cadre général avec Q quelconque et Σ quelconque, nous ne pouvons pas trouver un tel modèle. En revanche, il existe un modèle particulier dans lequel le skew est porté par cette seule composante V_t^{12} tandis que la volatilité est portée par les deux autres composantes V_t^{11} et V_t^{22} . La spécification des paramètres du modèle est la suivante :

$$QQ^\top = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 0 & \gamma \\ \gamma & 0 \end{pmatrix}.$$

La restriction à une matrice Q diagonale n'est pas vraiment une restriction à condition de considérer une matrice pleine M : un changement de base permet de diagonaliser la matrice Q sans perdre de généralité. On pourrait se poser la question de savoir si le facteur X_t apparaissant dans la dynamique de r_t est vraiment nécessaire. En fait, le facteur X_t permet de jouer sur la volatilité du taux court sans avoir d'influence réelle sur le skew.

En effet, dans ce modèle spécifié où le skew est porté par le facteur V_t^{12} , on peut calculer $Tr(\Sigma V_t) = 2\gamma V_t^{12}$. Si le facteur X_t n'était pas présent, le niveau du taux court $r_t = \alpha + Tr(\Sigma V_t)$ serait proportionnel au skew ce qui fournirait un modèle irréaliste et incompatible avec le marché. Ainsi, la présence du facteur X_t est fondamentale dans ce cadre en apportant la flexibilité nécessaire au modèle. La forme de X_t se justifie par le fait qu'on veut conserver une forme affine en donnant la forme la plus générale possible quitte à restreindre le modèle plus tard.

Étude du modèle pour des options sur taux court forward

Connaissant le prix du zéro-coupon par l'expression (3.1), on exprime le taux court forward en utilisant la relation :

$$f(t, T) = -\partial_T \ln B(t, T).$$

En utilisant les équation de Riccati et le fait que la structure soit homogène $A(t, T) = A(t - T)$, on obtient :

$$\begin{aligned} f(t, T) &= Tr[(\theta(t) - A(t, T)M - M^\top A(t, T) - 2A(t, T)Q^\top QA(t, T))V_t] \\ &\quad + [\delta(t) + P^\top H(t, T)]X_t - \beta Tr(Q^\top QA(t, T)) - H^\top(t, T)P\aleph \\ &\quad - \frac{1}{2}Tr[D_1 R^\top H(t, T)H^\top(t, T)R] + \alpha(t) \\ &= Tr[a(t, T)V_t] + b(t, T)X_t + c(t, T). \end{aligned}$$

Ainsi, en notant $b_k = \sum_{i=1}^p b_i(t, T)R^{ik}$, les expressions recherchées s'écrivent sous la forme :

$$\begin{aligned} d \langle f(., T) \rangle_t &= [2Tr(a(t, T)Q^\top Qa(t, T)^\top V_t) + \sum_{k=1}^h (\alpha_k + Tr(\Sigma_k V_t))b_k^2]dt, \\ d \langle f(., T), V^{ij} \rangle_t &= [(V_t a(t, T)Q^\top Q)^{ij} + (V_t a(t, T)Q^\top Q)^{ji}]dt. \end{aligned}$$

Les expression obtenues ont la même forme celles obtenues pour le taux court. Par conséquent, la condition pour obtenir un modèle dont le skew est représenté par un seul facteur peut s'écrire de la manière suivante :

$$QQ^\top = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \quad a(t, T) = \begin{pmatrix} 0 & \gamma(t, T) \\ \gamma(t, T) & 0 \end{pmatrix}.$$

Il faut trouver les conditions pour avoir $a(t, T)$ antidiagonale. Le fait d'avoir une matrice Σ antidiagonale ne suffit pas et il faut ajouter des conditions sur d'autres paramètres du modèle. Une condition suffisante est que les matrices Σ , Σ_j , Π_j soient antidiagonales et que la matrice M soit diagonale. En effet, le calcul de $a(t, T)$ est donné par l'équation de Riccati suivante :

$$\begin{aligned} a(t, T) = & \Sigma - \sum_{j=1}^p H(t, T)_{j1} \Pi_j - \frac{1}{2} \left(\sum_{j=1}^h (R^\top H(t, T) H^\top(t, T) R)^{jj} \Sigma_j \right) \\ & - \left(A(t, T) M + M^\top A(t, T) + 2A(t, T) Q^\top Q A(t, T) \right). \end{aligned}$$

Ainsi, il est possible d'assurer l'antidiagonalité de $a(t, T)$ et donc de $A(t, T)$ avec certaines conditions sur les paramètres. Par conséquent, on peut à nouveau spécifier un modèle dans lequel le skew d'une option sur taux court forward sera dirigé par le facteur V_t^{12} .

Ce modèle affine permet d'avoir une bonne idée de l'impact des facteurs et d'avoir une structure de volatilité assez riche. Il est possible de spécifier le modèle afin de contrôler la dynamique du skew. Nous allons à présent décrire l'évaluation des caplets et des swaptions qui sont indispensables pour la calibration du modèle.

Étude du modèle pour lequel le skew est un risque "hedgeable"

Nous proposons un modèle particulier très simple avec $n = 2$ et $p = h = 1$ dans lequel le skew pour une option sur taux court serait dirigé par le facteur V_t^{12} alors que la volatilité du taux court sera dirigé par les facteurs V_t^{11} et V_t^{22} .

$$\begin{cases} dX_t &= \lambda(\aleph - X_t)dt + Tr(\Pi V_t)dt + \sqrt{\alpha_1 + Tr(\Sigma_1 V_t)}dB_t \\ dV_t &= (\beta Q^\top Q + M V_t + V_t M^\top)dt + \sqrt{V_t}dW_t Q + Q^\top dW_t^\top \sqrt{V_t} \\ r_t &= l + \rho V_t^{12} + \sqrt{1 - \rho^2} X_t, \end{cases}$$

avec les restrictions d'avoir Q diagonale et Σ antidiagonale symétrique :

$$Q = \begin{pmatrix} Q^{11} & 0 \\ 0 & Q^{22} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 0 & \frac{\rho}{2} \\ \frac{\rho}{2} & 0 \end{pmatrix}.$$

Si on veut conserver un modèle dans lequel le taux court à une composante de retour à la moyenne, il faut imposer certaines restrictions sur les paramètres de la dynamique de X_t . Ces contraintes sont données par :

$$\Pi = -\frac{\rho}{2\sqrt{1-\rho^2}} \begin{pmatrix} 2m_{12} & \lambda + m_{11} + m_{22} \\ \lambda + m_{11} + m_{22} & 2m_{21} \end{pmatrix}, \quad \Sigma_1 = \begin{pmatrix} (Q^{jj})^2 & 0 \\ 0 & (Q^{ii})^2 \end{pmatrix}, \quad \alpha_1 = 0.$$

Un modèle équivalent sur le taux court avec retour à la moyenne peut s'écrire sous la forme suivante :

$$\left\{ \begin{array}{l} dr_t = \lambda(l - r_t)dt + \sqrt{(Q^{22})^2 V_t^{11} + (Q^{11})^2 V_t^{22}} dB_t \\ dV_t = (\beta Q^\top Q + M V_t + V_t M^\top)dt + \sqrt{V_t} dW_t Q + Q^\top dW_t^\top \sqrt{V_t} \\ < r_t, V_t^{12} > = \rho. \end{array} \right. \quad (3.8)$$

Dans ce modèle, la variance instantanée du taux court $< r_t >$ ne dépend que des facteurs V_t^{11} et V_t^{22} alors que la covariance entre le taux court et sa variance $< r_t, (Q^{22})^2 V_t^{11} + (Q^{11})^2 V_t^{22} >_t$ ne dépend que du facteur V_t^{12} . On considère la configuration suivante de paramètres permettant de retrouver le modèle (3.8) :

$$\begin{aligned} P = 0.1, \quad \Pi_1 &= \begin{pmatrix} 0 & -0.000502 \\ -0.000502 & 0 \end{pmatrix}, \quad \aleph = 0, \quad \Sigma_1 = \begin{pmatrix} 0.0001 & 0 \\ 0 & 0.0001 \end{pmatrix}, \quad \alpha_1 = 0, \\ Q &= \begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}, \quad M = \begin{pmatrix} -0.1 & 0 \\ 0 & -0.1 \end{pmatrix}, \quad \beta = 20, \\ \alpha = l = 0.04, \quad \Sigma &= \begin{pmatrix} 0 & -0.05 \\ -0.05 & 0 \end{pmatrix}, \quad \delta_1 = 0.995. \end{aligned}$$

L'augmentation du nombre de facteurs apporte une certaine flexibilité sur le smile de caplets et les surfaces de volatilité avec des formes plus variées. Les smiles générés ont des formes assez riches et on peut générer par exemple des formes convexes prononcées comme le montre le graphique suivant :

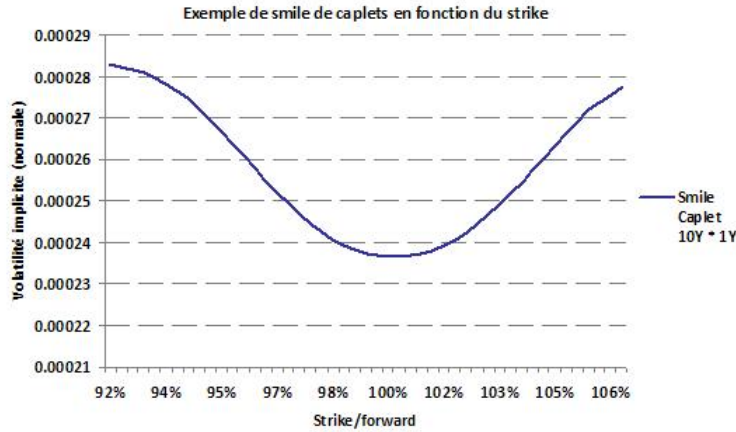


FIGURE 3.14 – Smile de caplets ($T = 10Y$, $U = 11Y$)

Concernant la surface de volatilité, elles sont aussi plus riches et il est d'ailleurs possible d'accentuer l'effet maturité avec des courbes qui s'écrasent rapidement lorsque la maturité augmente.

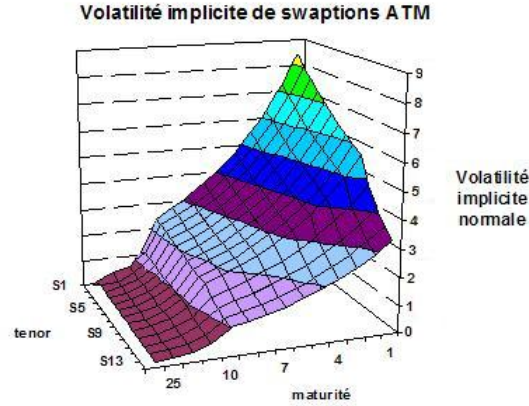


FIGURE 3.15 – Surface de volatilité ATM de prix de swaptions

Afin de faire une étude complète l'influence des différents paramètres du modèle, il faut analyser le contrôle de la dynamique du skew d'options sur taux court à l'aide du facteur V_t^{12} .

3.5 Étude du smile asymptotique

Dans cette partie, nous abordons les méthodes d'approximation de smile dans le cas où la maturité de l'option est petite devant l'échelle d'évolution de la volatilité. Il existe de nombreuses méthodes dans le monde des actions (voir Chapitre 2) mais il est plus difficile de mettre en place ces méthodes dans le monde des taux. Nous allons décrire des méthodes existantes et nous allons présenter la méthode des perturbations régulières adaptée au cadre du monde des taux d'intérêt.

3.5.1 Méthode d'approximation de smile en temps petit

Dans le monde des actions, il existe de nombreuses méthodes permettant d'obtenir des approximations de smiles. En particulier, il existe une méthode qui s'inspire des méthodes d'approximation en temps court introduites par V. Durrleman [11]. Cette méthode consiste à faire un développement de Taylor de la volatilité implicite en petites maturités et autour de la monnaie. En effet, en considérant la volatilité implicite $\Sigma_t(T, K)$ associé à un sous-jacent S_t , on peut écrire un développement de la volatilité implicite sous la forme :

$$\Sigma_t(T, K) \sim \Sigma_t(t, S_t) + (T - t)\Sigma'_T(t, S_t) + \frac{K - S_t}{S_t}\Sigma'_K(t, S_t) + \frac{1}{2}\left(\frac{K - S_t}{S_t}\right)\Sigma''_{KK}(t, S_t)$$

Cette méthode d'approximation en temps petit n'est pas directement applicable au modèle des taux avec une dépendance temporelle plus complexe. En effet, un sous-jacent $F_t(T)$

dans le monde des taux peut être martingale sous une probabilité forward neutre \mathbb{Q}_T de la manière suivante :

$$\begin{aligned}\frac{dF_t(T)}{F_t(T)} &= \vec{\sigma}(t, T)^\top d\vec{Y}_t^{\mathbb{Q}_T} \\ d\vec{\sigma}(t, T) &= \vec{a}_{1,t}(T)dt + \vec{a}_{2,t}(T)d\vec{Y}_t^{\mathbb{Q}_T} + \vec{a}_{3,t}(T)d\vec{Z}_t^{\mathbb{Q}_T}.\end{aligned}\quad (3.9)$$

La dynamique du sous-jacent $F_t(T)$ dépend fortement de la maturité T ce qui complique les méthodes d'approximation en temps petits. En effet, le fait de considérer que la maturité de l'option tend vers 0 modifie la dynamique du sous-jacent. Cependant, D. Nicolay a établi dans sa thèse [20] des formules en petites maturités et à la monnaie pour le skew, la convexité et de la pente sur une option dont le sous-jacent a la dynamique (3.9). À partir de ces notions, il est possible de déduire une approximation de la volatilité implicite en utilisant un développement de smile en strike et maturités.

Remarque 3.3. *L'étude de D. Nicolay révèle que la dépendance temporelle du sous-jacent $F_t(T)$ a un impact uniquement sur la pente Σ'_θ avec des termes correctifs additionnels. Il est possible d'appliquer ces travaux au modèle de Wishart afin d'obtenir des formules d'approximation de smile de caplets en courtes maturités.*

3.5.2 Méthode de perturbations régulières

La méthode des perturbations régulières introduites en finance par J.P. Fouque et al. [12, 13, 14] permet d'obtenir des approximations de prix d'options dans le monde des actions. Dans le chapitre 2, nous avons étudié l'adaptation de ces méthodes au cadre d'un modèle affine et nous avons constaté que la méthode consistant à perturber la transformée de Fourier du sous-jacent était efficace. Nous allons décrire la méthode des perturbations régulières appliquées aux modèles affines dans le monde des taux d'intérêt avec un développement spécifique au modèle de Wishart décrit dans ce chapitre. La structure des taux d'intérêt est plus complexe et il faut faire une approximation simultanée du prix des zéro-coupon et des prix d'options.

La méthode des perturbations régulières a l'avantage de ne pas considérer une maturité d'option qui tend vers 0 mais plutôt une maturité petite par rapport à l'échelle d'évolution de la volatilité. En effet, les options sur taux ont généralement des maturités plutôt longues et il apparaît intéressant de supposer plutôt que la maturité de l'option est petite par rapport aux temps caractéristiques d'évolution de la volatilité. Ainsi, l'idée est donc de perturber la dynamique de la volatilité de sorte que la maturité soit petite par rapport aux temps caractéristiques d'évolution de la volatilité.

La méthode de perturbation singulière est une perturbation "cohérente" des paramètres de la volatilité de sorte que la loi stationnaire du processus de volatilité soit non dégénérée.

Dans ce modèle, le processus de Wishart pilote la volatilité du taux court r_t et l'idée est donc de perturber les paramètres de la dynamique de V_t . Dans le chapitre 2, nous avons prouvé que la perturbation de la matrice M entraîne une contrainte sur la matrice Q de la forme :

$$M = \varepsilon \tilde{M} \quad Q = \sqrt{\varepsilon} \tilde{Q}$$

$$\tilde{M} = \begin{pmatrix} -m_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & -m_n \end{pmatrix} \quad \tilde{Q} = \begin{pmatrix} q_{11} & \dots & q_{1n} \\ \vdots & \ddots & \vdots \\ q_{n1} & \dots & q_{nn} \end{pmatrix}.$$

Nous considérons dans la suite que les paramètres M et Q sont perturbés de cette manière.

Développement du prix des zéro-coupon

Ayant supposé la perturbation des paramètres M et Q du processus de Wishart, nous allons chercher à établir une approximation du prix des zéro-coupon. Nous rappelons l'expression du prix des zéro-coupon qui est donnée par :

$$B(t, T) = \exp \left(Tr[A(t, T)V_t] + H^\top(t, T)X_t + C(t, T) \right),$$

avec les fonctions A , H et C vérifiant :

$$\left\{ \begin{array}{l} \partial_t A(t, T) + A(t, T)M + M^\top A(t, T) + 2A(t, T)Q^\top QA(t, T) \\ + \left(\sum_{j=1}^p H(t, T)^{j1} \Pi_j \right) + \frac{1}{2} \left(\sum_{j=1}^h (R^\top H(t, T)H^\top(t, T)R)^{jj} \Sigma_j \right) = \Sigma \\ A(T, T) = 0. \\ \\ H(t, T) = (P^\top)^{-1} \{ \exp[-P^\top(T-t)] - I_p \} \delta \\ H(T, T) = 0. \\ \\ \partial_t C(t, T) + \beta Tr[Q^\top QA(t, T)] + H^\top(t, T)P\aleph \\ + \frac{1}{2} Tr[D_0 R^\top H(t, T)H^\top(t, T)R] = \alpha \\ C(T, T) = 0. \end{array} \right.$$

La fonction H possède une forme explicite, alors que les fonctions A et C vérifient un système de Riccati. L'idée est de chercher un développement des fonctions A et C . La structure des équations de Riccati suggère de considérer un développement sous la forme :

$$A(t, T) = \sum_{i=0}^{+\infty} \epsilon^i A_i(t, T), \quad C(t, T) = \sum_{i=0}^{+\infty} \epsilon^i C_i(t, T).$$

Ainsi, en injectant les paramètres perturbés M et Q dans ces équations de Riccati, il est possible de déduire un développement des fonctions de Riccati A et C ce qui fournit un développement du prix des zéro-coupon. Un développement à l'ordre ϵ de ces fonctions est fait à l'annexe B.

Développement du prix de caplet

On suppose connu le développement des fonctions A et C . On rappelle que l'expression (3.2) du prix de caplet de maturité T sur un taux $L(T, T + \nu)$ est donné par :

$$\begin{aligned}\Psi(\gamma, t, T, \nu) &= \mathbb{E}_0^{\mathbb{Q}_{T+\nu}} [\exp \{ \gamma \log(1 + \nu L_t(T, T + \nu)) \}] \\ &= \exp \{ \gamma (C(t, T) - C(t, T + \nu)) \} \\ &\quad \Phi(0, t, T + \nu, \gamma(A(t, T) - A(t, T + \nu)), \gamma(H(t, T) - H(t, T + \nu)), V_0, X_0).\end{aligned}$$

Ainsi, un développement de la transformée des facteurs Φ permet d'obtenir un développement du prix de caplet. La transformée des facteurs Φ s'exprime sous la forme :

$$\Phi(t, T, U, \Gamma, \Lambda, V_t, X_t) = \exp \left(\text{Tr}[D(t, T, U, \Gamma, \Lambda)V_t] + [J(t, T, U, \Gamma, \Lambda)]^\top X_t + Z(t, T, U, \Gamma, \Lambda) \right),$$

où les fonctions D , J et Z vérifient le système :

$$\left\{ \begin{array}{l} \partial_t D + D[M'(t, U)] + [M'(t, U)]^\top D + 2D[Q^\top Q]D \\ + \left(\sum_{j=1}^h [(R^\top H(t, U)JR)^{jj} + \frac{1}{2}(R^\top JJ^\top R)^{jj}] \Sigma_j \right) \\ + \left(\sum_{j=1}^p J(t, T, U, \Gamma, \Lambda)^{j1} \Pi_j \right) = 0 \\ D(T, T, U, \Gamma, \Lambda) = \Gamma. \\ \\ J(t, T, U, \Gamma, \Lambda) = \exp[-P^\top(T - t)]\Lambda \\ J(T, T, U, \Gamma, \Lambda) = \Lambda. \\ \\ \partial_t Z(t, T) + \beta \text{Tr}[Q^\top QD] + J^\top P\aleph + \text{Tr}[D_0 \{ \frac{1}{2}R^\top JJ^\top R + R^\top H(t, U)J^\top R \}] = 0 \\ Z(T, T, U, \Gamma, \Lambda) = 0. \end{array} \right.$$

L'expression du prix de caplet étant donnée par

$$\begin{aligned}\Psi(\gamma, t, T, \nu) &= \exp \{ \gamma (C(t, T) - C(t, T + \nu)) \} \\ &\quad \Phi(0, t, T + \nu, \gamma(A(t, T) - A(t, T + \nu)), \gamma(H(t, T) - H(t, T + \nu)), V_0, X_0).\end{aligned}$$

On cherche une expression des valeurs terminales Λ et Γ des fonctions D et J sous la forme :

$$\begin{aligned}\Gamma &= \gamma[A(t, T) - A(t, T + \nu)] = \gamma \sum_{i=1}^{+\infty} [A_i(T - t) - A_i(T + \nu - t)]. \\ \Lambda &= \gamma[H(t, T) - H(t, T + \nu)] = \gamma(P^\top)^{-1} e^{-(T-t)P^\top} (I_p - e^{P^\top \nu}) \delta.\end{aligned}$$

L'expression de la fonction J est explicite :

$$J(t, T, U, \Gamma, \Lambda) = \gamma(P^\top)^{-1} e^{-2(T-t)P^\top} (I_p - e^{-P^\top(U-T)}) \delta.$$

On cherche un développement des fonctions D et Z sous la forme :

$$D(t, T) = \sum_{i=0}^{+\infty} \epsilon^i D_i(t, T), \quad Z(t, T) = \sum_{i=0}^{+\infty} \epsilon^i Z_i(t, T).$$

En revanche, il faut effectuer un travail particulier pour trouver le développement de la fonction D . En effet, il faut perturber l'équation de Riccati vérifiée par la fonction D dont la condition terminale $D(T, T, U, \Gamma, \Lambda) = \gamma[A(t, T) - A(t, U)]$ dépend de la fonction A qui est elle-même solution d'un système de Riccati. L'idée est d'effectuer un développement de la fonction A , de l'injecter dans l'équation et de déduire un développement de la fonction D . D'ailleurs, si les fonctions A et C possèdent un développement à l'ordre ϵ , il est possible d'avoir un développement à l'ordre ϵ des fonctions D et Z ce qui fournit une approximation à l'ordre ϵ du prix de caplet. Un développement des fonctions D et Z à l'ordre ϵ est donnée à l'annexe C.

Il est donc possible d'adapter la méthode de perturbation singulière au cadre des modèles affines sur les taux d'intérêt. Cela nécessite un travail important dans la mesure où il faut perturber deux systèmes de Riccati imbriqués avec des paramètres dépendant du temps. Il est possible de faire le même type de développements pour les swaptions.

3.6 Conclusion

Nous avons étudié un modèle affine dans le monde des taux d'intérêt qui possède des propriétés intéressantes avec des structures de volatilité et de corrélation assez riches. Ce modèle est une généralisation multidimensionnelle du modèle de Heston qui s'inspire du modèle de Wishart dans le monde des actions et qui est adapté à la modélisation des taux d'intérêt.

Nous avons détaillé des méthodes d'évaluation des produits vanilles comme les caplets et les swaptions. La méthode des moments spécifique aux modèles affines est efficace et est facilement adaptable à notre cadre. Nous avons aussi présenté une étude rapide de l'influence des paramètres du modèle de Heston sur les taux afin de comprendre l'impact des différents types de paramètres.

Ce modèle est assez intéressant dans la mesure où la structure de volatilité est plus complexe que celle modèle affine standard dans la spécification de Daï Singleton. Nous définissons une projection du modèle sur l'espace des modèles affines standards et il serait intéressant de comparer le comportement de ces deux modèles. De plus, nous introduisons une spécification des paramètres de manière à contrôler la corrélation du taux court de sa volatilité par un facteur indépendant de la volatilité.

Enfin, nous avons adapté la méthode des perturbations régulières à ce modèle qui nécessite la perturbation simultanée du système de Riccati du prix des zéro-coupon et des options.

Il reste de nombreuses études théoriques et numériques à effectuer sur ce type de modèle qui pourrait avoir des propriétés très intéressantes. Il faudrait aussi considérer des méthodes d'évaluation des options exotiques avec une méthode de Monte Carlo. Il existe de nombreux travaux récents sur la simulation efficace du processus de Wishart [1] et il serait intéressant de l'appliquer à ce modèle pour l'évaluation d'options exotiques.

3.7 Annexes

Annexe A : Outils mathématiques

Annexe A1 : Résolution des Riccati par linéarisation

On considère les équations de Riccati de la forme suivante :

$$\begin{cases} A'(t) + A(t)QA(t) + A(t)M + NA(t) + \Gamma(t) = 0 \\ A(T) = \Psi \end{cases}$$

La méthode de linéarisation consiste à chercher la solution sous la forme suivante :

$$\begin{cases} A(t) = [F(t)]^{-1}G(t) \\ F(T) = I_n \\ G(T) = \Psi. \end{cases}$$

En injectant cette expression dans l'équation et en identifiant les termes qui contiennent A et ceux qui ne contiennent pas A , on obtient :

$$\begin{cases} -F'(t) + F(t)N + G(t)Q = 0 \\ G'(t) + F(t)\Gamma(t) + G(t)M = 0 \\ F(T) = I_n \\ G(T) = \Psi. \end{cases}$$

Le système peut s'écrire de la manière suivante :

$$\begin{cases} [F'(t) \ G'(t)] = [F(t) \ G(t)] \begin{pmatrix} N & -\Gamma(t) \\ Q & -M \end{pmatrix} \\ F(T) = I_n \\ G(T) = \Psi. \end{cases}$$

On a ainsi obtenu une expression quasi fermée qu'il est possible de résoudre par des méthodes numériques utilisées pour traiter les équations différentielles (type Runge-Kutta).

Lorsque la fonction Γ ne dépend pas du temps, on peut obtenir une expression explicite en utilisant l'exponentielle matricielle.

$$\left\{ \begin{array}{l} A(t) = [F(t)]^{-1}G(t) \\ [F(t) \ G(t)] = [I_n \ \Psi] \exp \left\{ \begin{pmatrix} -(T-t)N & (T-t)\Gamma \\ -(T-t)Q & (T-t)M \end{pmatrix} \right\} \end{array} \right\}.$$

Annexe A.2 : Calcul explicite de $\int_t^T \exp[(u-s)M]Q \exp[(u-r)N]du = \int_t^T R(t)dt$

On cherche à calculer (sans utiliser de méthodes numériques) la fonction suivante :

$$U(t) = \int_t^T \exp[(u-s)M]Q \exp[(u-r)N]du = \int_t^T R(t)dt.$$

La fonction $R(t)$ est solution de l'équation de Riccati suivante :

$$\left\{ \begin{array}{l} R'(t) + Q + R(t)N + MR(t) = 0 \\ R(T) = 0. \end{array} \right.$$

En utilisant l'annexe A.1, on peut obtenir la solution sous la forme suivante :

$$\left\{ \begin{array}{l} U(t) = \exp[(t-s)M][F(t)]^{-1}G(t) \exp[(t-r)N] \\ [F(t) \ G(t)] = [I_n \ 0] \exp \left\{ \begin{pmatrix} -(T-t)M & (T-t)Q \\ 0 & (T-t)N \end{pmatrix} \right\} \end{array} \right\}.$$

Annexe A.3 : Calcul explicite de $\partial_{M_{11}}U(t, T, r, s, M, N, \Omega)$

En utilisant l'expression de la fonction U , on sait que :

$$\partial_{M_{11}}U(t, T, r, s, M, N, \Omega) = \partial_{M_{11}}e^{(T-s)M}F(t)^{-1}G(t)e^{(t-r)N}.$$

Or la fonction G ne dépend pas de M et la fonction F vérifie l'équation $F'(t) = F(t)M$ avec pour condition terminale $F(T) = I_n$. Cela fournit $F(t) = e^{M(T-t)}$ ce qui donne :

$$\begin{aligned} \partial_{M_{11}}U(t, T, r, s, M, N, \Omega) &= (t-s)U(t, T, r, s, M, N, \Omega) + (T-t)U(t, T, r, s, M, N, \Omega) \\ &= (T-s)U(t, T, r, s, M, N, \Omega). \end{aligned}$$

Annexe B : Développement du prix des zéro-coupon

On s'intéresse à une approximation du prix des zéro-coupon à l'ordre ϵ . Pour cela, il suffit de trouver des développement des fonctions A et C du système de Riccati.

Développement de la fonction A

L'équation de Riccati vérifiée par A est donnée par :

$$\begin{aligned} \partial_t A(t, T) + A(t, T)M + M^\top A(t, T) + 2A(t, T)Q^\top QA(t, T) \\ = \Sigma - \left(\sum_{j=1}^p H(t, T)^{j1} \Pi_j \right) - \frac{1}{2} \left(\sum_{j=1}^h (R^\top H(t, T) H^\top(t, T) R)^{jj} \Sigma_j \right) \\ = \Theta_A(t, T). \end{aligned}$$

Ordre 0 : À l'ordre 0, on a $\partial_t A_0(t, T) = \Theta_A(t, T)$, en intégrant entre t et T , on obtient :

$$\begin{aligned} A_0(t, T) &= - \int_t^T \Theta_A(s, T) ds \\ &= -\Sigma(T-t) + \sum_{j=1}^h \left(\int_t^T H(s, T)^{j1} ds \right) \Pi_j + \frac{1}{2} \sum_{j=1}^h \left(\int_t^T (R^\top H(s, T) H^\top(s, T) R)^{jj} ds \right) \Sigma_j. \end{aligned}$$

- Le calcul de $\int_t^T H(s, T) ds$ donne immédiatement :

$$\int_t^T H(s, T) ds = -(P^\top)^{-1} [H(t, T) + (T-t)\delta].$$

- Calcul de $\int_t^T H(s, T) H^\top(s, T) ds$:

On peut écrire que :

$$\begin{aligned} P^\top \int_t^T H(s, T) H^\top(s, T) ds P &= \int_t^T e^{-P^\top(T-s)} \delta \delta^\top e^{-P(T-s)} + (T-t) \delta \delta^\top \\ &+ (P^\top)^{-1} [e^{-P^\top(T-t)} - I_p] \delta \delta^\top + \delta \delta^\top [e^{-P(T-t)} - I_p] P^{-1}. \end{aligned}$$

En considérant l'annexe A.2 donnant l'expression de la fonction U , on obtient :

$$\begin{aligned} \int_t^T H(s, T) H^\top(s, T) ds &= (P^\top)^{-1} U(t, T, T, T, P^\top, P, \delta \delta^\top) P^{-1} + (T-t) (P^\top)^{-1} \delta \delta^\top P^{-1} \\ &+ (P^\top)^{-2} [e^{-P^\top(T-t)} - I_p] \delta \delta^\top P^{-1} + (P^\top)^{-1} \delta \delta^\top [e^{-P(T-t)} - I_p] P^{-2}. \end{aligned}$$

Ainsi, l'expression de A_0 est obtenu avec le calcul explicite de ces intégrales.

Ordre ϵ : En analysant les termes d'ordre ϵ de l'équation vérifiée par A , on a :

$$\partial_t A_1(t, T) + A_0(t, T) \tilde{M} + \tilde{M}^\top A_0(t, T) + 2A_0(t, T) \tilde{Q}^\top \tilde{Q} A_0(t, T) = 0.$$

Ainsi, on déduit l'expression de A_1 sous la forme :

$$A_1(t, T) = \left(\int_t^T A_0(s, T) ds \right) M + M^\top \left(\int_t^T A_0(s, T) ds \right) + 2 \left(\int_t^T A_0(s, T) Q^\top Q A_0(s, T) ds \right).$$

On pourrait calculer ces intégrales numériquement puisque la fonction A_0 est connue explicitement. On détaille pour le lecteur quelques astuces permettant de ne pas avoir à effectuer de résolution numérique de ces intégrales.

- Calcul de $\int_t^T A_0(s, T) ds$:
En utilisant l'expression de A_0 , on obtient :

$$\begin{aligned} \int_t^T A_0(s, T) ds &= -\Sigma \int_t^T (T-s) ds + \sum_{j=1}^h \int_t^T \left(\int_s^T H(u, T)^{j1} du \right) ds \Pi_j \\ &\quad + \frac{1}{2} \sum_{j=1}^h \left[R^\top \left(\int_t^T \left(\int_s^T H(u, T) H^\top(u, T) du \right) ds \right) R \right]^{jj} \Sigma_j. \end{aligned}$$

Ainsi, il faut calculer chacune des intégrales de cette expression :

- Un calcul rapide fournit l'expression :

$$\int_t^T \left(\int_s^T H(u, T) du \right) ds = (P^\top)^{-2} H(t, T) + (T-t)(P^\top)^{-2} \delta + \frac{(T-t)^2}{2} (P^\top)^{-1} \delta.$$

- Concernant $\int_t^T \left(\int_s^T H(u, T) H^\top(u, T) du \right) ds$, on a :

$$\begin{aligned} P^\top \int_t^T \left(\int_s^T H(u, T) H^\top(u, T) du \right) ds P &= \int_t^T \int_s^T e^{-P^\top(T-u)} \delta \delta^\top e^{-P(T-u)} du ds \\ &\quad + \int_t^T \int_s^T \delta \delta^\top I_p - \delta \delta^\top \int_t^T \int_s^T e^{-P(T-u)} du ds \\ &\quad - \int_t^T \int_s^T e^{-P^\top(T-u)} du ds \delta \delta^\top. \end{aligned}$$

La seule intégrale complexe à calculer est le premier terme mais il est possible d'utiliser un théorème de Fubini ainsi qu'une astuce intéressante utilisant l'annexe A.2. En effet, on peut écrire :

$$\begin{aligned} \int_t^T \int_s^T e^{-P^\top(T-u)} \delta \delta^\top e^{-P(T-u)} du ds &= \int_t^T (u-t) e^{-P^\top(T-u)} \delta \delta^\top e^{-P(T-u)} du \\ &= (T-t) U(t, T, T, TP^\top, P, \delta \delta^\top) \\ &\quad + \partial_{M_{11}} U(t, T, T, TP^\top, P, \delta \delta^\top), \end{aligned}$$

en ayant utilisé la propriété suivante qui est prouvée dans l'annexe A.3

$$\partial_{M_{11}} U(t, T, r, sM, N, \Omega) = \int_t^T (u-s) e^{(u-s)M} \Omega e^{(u-r)N} du.$$

On en déduit finalement l'expression de cette intégrale

$$\begin{aligned}
 \int_t^T \left(\int_s^T H(u, T) H^\top(u, T) du \right) ds &= (P^\top)^{-1} (T - t) U(t, T, T, TP^\top, P, \delta \delta^\top) P^{-1} \\
 &+ (P^\top)^{-1} \partial_{M_{11}} U(t, T, T, TP^\top, P, \delta \delta^\top) P^{-1} \\
 &- (P^\top)^{-1} \delta \delta^\top [P^{-2} (T - t) - P^{-3} (I_p - e^{-P(T-t)})] \\
 &- [(P^\top)^{-2} (T - t) - (P^\top)^{-3} (I_p - e^{-P(T-t)})] \delta \delta^\top P^{-1} \\
 &+ \frac{(T - t)^2}{2} \delta \delta^\top I_p.
 \end{aligned}$$

Ainsi, il est possible de trouver des astuces intéressantes afin d'éviter une résolution numérique. Ces astuces deviennent utiles si on envisage de pousser un développement à un ordre supérieur comme ϵ^2 .

Développement de la fonction C

L'équation de Riccati vérifiée par la fonction C se réécrit :

$$\begin{aligned}
 \partial_t C(t, T) + \beta \text{Tr}[Q^\top Q A(t, T)] &= \alpha - H^\top(t, T) P \aleph - \frac{1}{2} \text{Tr}[D_0 R^\top H(t, T) H^\top(t, T) R] \\
 &= \Theta_C(t, T)
 \end{aligned}$$

Ordre 0 :

En intégrant entre t et T , on obtient :

$$\begin{aligned}
 C_0(t, T) &= - \int_t^T \Theta_C(t, T) \\
 &= -\alpha(T - t) - (T - t) \delta^T \aleph - H^\top(t, T) \aleph \\
 &\quad + \frac{1}{2} \text{Tr} \left[D_0 R^\top \left(\int_t^T H(s, T) H^\top(s, T) ds \right) R \right]
 \end{aligned}$$

dont le calcul a déjà été explicité précédemment.

Ordre ϵ :

En identifiant l'équation et en l'intégrant entre t et T , on obtient :

$$C_1(t, T) = -\beta \text{Tr} \left[\tilde{Q}^\top \tilde{Q} \left(\int_t^T A_0(s, T) ds \right) \right],$$

avec le calcul de l'intégrale $\int_t^T A_0(s, T) ds$ qui a été détaillé précédemment.

Annexe C : Développement du prix de caplet

On cherche à établir une approximation du prix de caplet à l'ordre ϵ et pour cela on considère un développement de la transformée des facteurs Φ . On sait déjà que la fonction J possède une expression explicite donnée par :

$$J(t, T, U) = \gamma(P^\top)^{-1} e^{-2(T-t)P^\top} (I_p - e^{-P^\top(U-T)}) \delta.$$

Développement de la fonction D

On réécrit l'équation vérifiée par la fonction D

$$\begin{aligned} \partial_t D(t, T, U) + D(t, T, U)[M'(t, U)] + [M'(t, U)]^\top D(t, T, U) + 2D(t, T, U)Q^\top QD(t, T, U) \\ = - \sum_{j=1}^h [(R^\top H(t, U)J^\top(t, T, U)R)^{jj} + \frac{1}{2}(R^\top J(t, T, U)J^\top(t, T, U)R)^{jj}] \Sigma_j \\ - \sum_{j=1}^p J(t, T, U, \Gamma, \Lambda)^{j1} \Pi_j = 0 \\ = \Theta_D(t, T). \end{aligned}$$

De plus, la fonction D vérifie $D(T, T, U, \Gamma, \Lambda) = \Gamma$ et on sait que Γ possède aussi un développement sous la forme :

$$\Gamma = \sum_{i=0}^{+\infty} \epsilon^i \Gamma_i(t, T) = \sum_{i=0}^{+\infty} \epsilon^i (A_i(t, T) - A_i(t, U)).$$

On cherche le développement de D à l'ordre ϵ .

Ordre 0 :

$$\begin{aligned} D_0(t, T, U) &= \Gamma_0(t, T) - \int_t^T \theta_I(s, T) ds \\ &= A_0(t, T) - A_0(t, U) + \sum_{j=1}^h \left[R^\top \left(\int_t^T H(s, U) J^\top(s, T, U) ds \right) R \right]^{jj} \Sigma_j \\ &\quad + \frac{1}{2} \sum_{j=1}^h \left[R^\top \left(J(s, T, U) J^\top(s, T, U) ds \right) R \right]^{jj} \Sigma_j \\ &\quad + \sum_{j=1}^p \left(\int_t^T J(s, T, U, \Gamma, \Lambda) ds \right)^{j1} \Pi_j. \end{aligned}$$

Ainsi, le calcul de D_0 se résume au calcul de plusieurs intégrales.

- Calcul de $\int_t^T H(s, U)J^\top(s, T, U)ds$:

$$\begin{aligned} \int_t^T H(s, U)J^\top(s, T, U)ds &= -\frac{1}{2}\gamma(P^\top)^{-1}\delta\delta^\top(I_p - e^{-P(U-T)})P^{-2}(I_p - e^{-P(T-t)}) \\ &\quad + \gamma(P^\top)^{-1}U\left(t, T, T, U, P^\top, 2P, \delta\delta^\top(I_p - e^{-P(T-U)})\right)P^{-1}. \end{aligned}$$

- Calcul de $\int_t^T J(s, T, U)J^\top(s, T, U)ds$:

$$\int_t^T J(s, T, U)J^\top(s, T, U)ds = \gamma^2(P^\top)^{-1}U\left(t, T, T, T, 2P^\top, 2P, (I_p - e^{-P^\top(T-U)})\delta\delta^\top(I_p - e^{-P(T-U)})\right)P^{-1}.$$

- La troisième intégrale à calculer est $\int_t^T J(s, T, U)ds$:

$$\int_t^T J(s, T, U)ds = \frac{1}{2}\gamma(P^\top)^{-2}[I_p - e^{-2P^\top(T-t)}][I_p - e^{-P^\top(T-U)}].$$

Ordre ϵ :

$$\begin{aligned} D_1(t, T, U) &= A_1(t, T) - A_1(t, U) + \int_t^T D_0(s, T, U)\left(\tilde{M} + 2\tilde{Q}^\top\tilde{Q}A_0(s, U)\right)ds \\ &\quad + \int_t^t (\tilde{M}^\top + 2A_0(t, U)\tilde{Q}^\top\tilde{Q})D_0(s, T, U)ds \\ &\quad + 2\int_t^T D_0(s, T, U)\tilde{Q}^\top\tilde{Q}D_0(s, T, U)ds. \end{aligned}$$

Pour calculer l'ensemble de ces intégrales, on peut essayer de faire des développements pour avoir une expression explicite ou on peut faire un calcul numérique rapide puisque les fonctions A_0 et I_0 ont une expression explicite. L'idéal serait d'avoir des expressions explicites afin d'avoir un développement à l'ordre ϵ^2 . Il existe de nombreuses astuces permettant d'obtenir des expressions explicites de la plupart de ces intégrales.

Développement de la fonction Z

On peut aussi expliciter un développement de Z à l'ordre ϵ . L'équation de Riccati qui lui est associée se réécrit :

$$\begin{aligned} \partial_t Z(t, T, U) + \beta \text{Tr}[Q^\top Q D(t, T, U)] &= -J^\top(t, T, U)P\aleph - \frac{1}{2}\text{Tr}[D_0R^\top J(t, T, U)J^\top(t, T, U)R] \\ &\quad - \text{Tr}[D_0R^\top H(t, U)J^\top(t, T, U)R] \\ &= \Theta_Z(t, T). \end{aligned}$$

Ordre 0 :

$$\begin{aligned}
 Z_0(t, T, U) &= - \int_t^T \Theta_Z(s, T) ds \\
 &= \int_t^T J^\top(s, T, U) P \mathbb{N} ds + \frac{1}{2} Tr \left[D_0 R^\top \left(\int_t^T J(s, T, U) J^\top(s, T, U) ds \right) R \right] \\
 &\quad + Tr \left[D_0 R^\top \left(\int_t^T H(s, U) J^\top(s, T, U) ds \right) R \right]
 \end{aligned}$$

dont toutes les intégrales sont calculables explicitement.

Ordre ϵ :

$$Z_1(t, T, U) = \beta Tr \left[Q^\top Q \left(\int_t^T D_0(s, T, U) ds \right) \right]$$

dont le calcul a déjà été donné explicitement.

Bibliographie

- [1] Abdelkoddousse A. and A. Alfonsi (2010), “Exact and high order discretization schemes for Wishart processes and their affine extensions”, *Working Paper*, [http ://hal.archives-ouvertes.fr/docs/00/49/13/71/PDF/](http://hal.archives-ouvertes.fr/docs/00/49/13/71/PDF/).
- [2] Bru M.F. (1991), “Wishart Processes”, *Journal of Theoretical Probability*, 4, 724-743.
- [3] Carr P. and D. Madan (1999), “Option valuation using the fast Fourier transform”, *Journal of Computational Finance*, 2, 61-73.
- [4] Collin-Dufresne P. and R.S. Goldstein (2002), “Pricing Swaptions Within an Affine Framework”, *The Journal of derivatives*, 9-26.
- [5] Da Fonseca J., M. Grasselli and C. Tebaldi (2005), “Wishart Multi-dimensional stochastic volatility”, *Working paper ESILV RR-31*.
- [6] Da Fonseca J., M. Grasselli and C. Tebaldi (2007), “Option pricing when correlations are stochastic : an analytical framework”, *Review of Derivatives Research*, 10, 2, 151-180.
- [7] Da Fonseca J., M. Grasselli and C. Tebaldi (2008), “A Multifactor Volatility Heston Model”, *Quantitative Finance*, 8, 6, 591-604.
- [8] Dai Q. and K.J. Singleton (2000), “Specification Analysis of Affine Term Structure Models”, *The Journal of Finance*, LV, 5, 1943-1978.
- [9] Duffie D., D. Filipovic and W. Schachermayer (2003), “Affine processes and applications in finance”, *Annals of Applied Probability*, 13, 984-1053.
- [10] Duffie G.R. (2002), “Term Premia and Interest Rate Forecasts in Affine Models”, *Journal of Finance*, 57, 405-443.
- [11] Durrleman V. (2004), “From Implied to Spot Volatilities”, Dissertation, *Princeton university*, [http ://www.cmap.polytechnique.fr/%7Evaldo/papers/dissertation.pdf](http://www.cmap.polytechnique.fr/%7Evaldo/papers/dissertation.pdf).
- [12] Fouque J.P., G. Papanicolaou, R. Sircar and K. Solna (2003), “Singular Perturbations in Option Pricing”, *SIAM Journal on Applied Mathematics*, 63, 5, 1648-1665.
- [13] Fouque J.P., G. Papanicolaou, R. Sircar and K. Solna (2003), “Multiscale stochastic volatility asymptotics”, *SIAM Journal Multiscale Modeling and Simulation*, 2, 1, 22-42.
- [14] Fouque J.P. and C.H. Han (2005), “Evaluation of Compound Options using Perturbation Approximation”, *Journal of Computational Finance*, 9, 1.

BIBLIOGRAPHIE

- [15] Gouriéroux C. (2006), “Wishart Processes for Stochastic Risk”, *Econometric Reviews, special issue on Stochastic Volatility*, 25, 1-41.
- [16] Gouriéroux C. and R. Sufana (2007), “Derivative Pricing with Multivariate Stochastic Volatility : Application to Credit Risk”, *Working Paper CREF 04-09*, [http ://www.yorku.ca/rsufana/dpmsv.pdf](http://www.yorku.ca/rsufana/dpmsv.pdf).
- [17] Hall P. (1992), “The bootstrap and Edgeworth expansion”, *New York Springer*.
- [18] Heath D., R. Jarrow, and A. Morton (1992), “Bond Pricing and the Term Structure of Interest Rates : A New Methodology for Contingent Claims Evaluation”, *Econometrica*, 60, 77-105.
- [19] Hull, J. and A. White (1990) “Pricing Interest Rate Derivative Securities”, *The Review of Financial Studies*, 3, 4, 573-592.
- [20] Nicolay D. (2010) “Volatility Dynamics”, Dissertation, in preparation.

Chapitre 4

Understanding, Modelling and Managing Longevity Risk

Introduction

This chapter is inspired by an article [6] published in Scandinavian Journal. This survey allows to understand the notion of longevity risk and provides the challenges concerning the management of this long term risk.

Motivations

The sustained improvements in longevity currently observed, are producing a considerable number of new issues and challenges at multiple societal levels. This is causing pensions to be one of the most highly publicised industries to be impacted by rising longevity. In 2009 many companies based in developed countries had closed the defined benefit retirement plans previously offered to their employees (such as the 401(K) pension plans in the United States). Broadly speaking, such scheme trends indicate a risk transfer from both the industry and the insurers, back to the policyholders, which from a social point of view is no longer satisfactory. Similarly, in several countries, defined benefit pension plans have continued to be replaced with defined contribution plans, which will most likely lead to the same result. Furthermore, a number of governments are set to increase retirement ages by 2 to 5 years to be able to take into account the changing dynamics of longevity improvements, and the impacts of ageing populations upon the financing of pensions.

The insurance industry is also facing a number of specific challenges related to longevity risk (i.e. the risk that trends in longevity improvements will change significantly in the future). In practical terms, greater capital has to be constituted to balance the long-term risk, a requirement only reinforced by the impacts of the financial crisis and new European regulation. Consequently, it has become paramount for insurance companies and pension

funds to find a suitable and efficient way to cross-hedge, or to transfer part of the longevity risk to reinsurers and the financial markets. However, in the absence of both theoretical consensus and established industry practices, the transfer of longevity risk is a difficult process to understand, and therefore, manage. In particular, because of the long-term nature of the risks, accurate longevity projections are delicate, and modelling the embedded interest rate risks remains challenging.

Prospective life tables containing longevity trend projections are frequently used to better manage longevity risk, proving to be particularly effective in reserving in life insurance. However, the irregularity of table updates can cause significant problems. For example, consider the French prospective life tables; they were updated in 2006, replacing the previous set of tables from 1993. The resulting disparities between the 1993 prospective tables and observed longevity caused French insurers to sharply increase their reserves by an average of 8%. In addition, insurance companies and pension funds are facing basis risk as the evolution of policyholder mortality is thought to differ from that of the national population as a result of selection effects. These selectivity effects may have different impacts on different portfolios, since the pace of change and levels of mortality are highly heterogeneous in the insurance industry, which makes it hard for insurance companies and pension funds to rely on national (or even industry) indices to manage their own longevity risk.

Presentation of the paper

To better understand longevity risk, its dynamics, causes, and the aforementioned heterogeneity of longevity improvements must be studied carefully, and aside from shorter-term managerial oscillations concerning average trends. Among the many standard stochastic models for mortality, a number have been inspired by the classical credit risk and interest rate literature, and as such produce a limited definition of mortality by age and time. An alternative is the microscopic modelling approach, which can be used for populations where individuals are characterised not only by age, but also by additional indicators that are reflective of lifestyle and living conditions. Such models can provide useful benefits for the risk analysis of a given insurance portfolio. Furthermore, when combined with studies on demographic rates, such as fertility and immigration, microscopic modelling is highly applicable at social and political levels, offering guidance for governmental strategies concerning, for example, immigration and retirement age policy.

The need to carry out microscopic studies is all the more apparent when the size of the considered portfolio is small and in all likelihood highly heterogeneous. This has proven particularly true when looking at the life settlement market (i.e. when owners of a life insurance policy sell an unneeded policy to a third party for more than its cash value and less than its face value). Although this market phenomenon carries anecdotal significance, it will not be studied specifically here. Rather, this paper focuses on longevity risk for the

larger portfolio, often referred to as macro-longevity risk.

The European insurance industry will soon be expected to comply to the new Solvency II Directive. The regulatory standards of the directive place emphasis on the manner in which insurance company endorsed risk should be handled so that it can withstand adverse economic and demographic situations. The regulation is scheduled to come into effect by late 2012, and will certainly enhance the development of alternative risk transfer solutions for insurance-risk generally, as well as for longevity-risk in particular. There is little doubt that the pricing methodologies for insurance related transactions, and specifically longevity-linked securities will be impacted as greater numbers of alternative solutions appear in the market. Today, the longevity market is immature and incomplete, with an evident lack of liquidity. Standard replication strategies are impossible, making the classical financial methodology unapplicable. In this case, indifference pricing, involving utility maximisation, seems to be a more appropriate strategy. In any case, due to the long maturities of the underlying risk, the modelling of long term interest rates is unavoidable, adding to the complexity of the problem.

The following paper is organised as follows : Section 1 provides an outline of the main characteristics of longevity risk, focusing on the classical and prospective life tables, mortality data, and specific features such as the cohort effect and basis risk. Section 2 presents the key models for mortality-risk, discussing how they can be used to model longevity-risk. Section 3 is dedicated to longevity risk management in reference to the new solvency regulations that stand to be enforced by 2012. Section 4 centres on longevity risk transfer issues, and the convergence between the insurance industry and capital markets. The final section looks at a number of key modelling questions concerning the pricing of longevity risk.

4.1 Main characteristics of longevity risk

In this section, the main characteristics of longevity risk are outlined according to available data and important longevity and mortality risk related issues.

4.1.1 Age, period and cohort life tables and available mortality data

The analysis of mortality of a given population or insured portfolio can take multiple forms, depending on the purposes of the study, the data available, and its reliability. For such analysis, the period life and cohort life tables have proven to be most useful. The Lexis diagram in Figure 4.1 is an age-period-cohort diagram ; the x-axis features calendar years and the y-axis corresponds to age. This representation enables the statistical understanding of both cohort and period life tables. Cohort life tables may be obtained from the diagonals of the Lexis diagram ; period life tables are given by vertical bands. Both continuous and discrete-time analysis are feasible, the latter being most applied to national data, since the

mortality of national populations is generally described by yearly observations¹. The equivalence between discrete and continuous methods is based upon the assumption that the mortality force (i.e. instantaneous mortality hazard rate) is constant (or does not fluctuate too much) within each square on the Lexis diagram.

However, the assumption of constant yearly mortality has to be avoided when observations of mortality are deeper than at a national level. Life insurers typically collect data for insured portfolios that are more accurate and that gives the actual ages attained by each insured individual. Thus, actuaries are used to derive the underlying mortality of such portfolios, taking into account censored observations, using the Kaplan-Meier estimator (see [36]). It is therefore possible to build time-continuous life tables.

The difference between national mortality data and that of an insured portfolio is not only limited to the continuity of observations. First of all, one key difference between national mortality and specific mortality in some insured portfolios, is the range of the observation period. The period for which observed mortality rates for the insurance portfolio are available is usually limited, often in the range of only 5 to 15 years. In contrast, national data can range from 100 to 200 years (see the Human Mortality Database (HMD) for example). This is one of the main reasons why in order to determine the actual level and the future trend of mortality, actuaries and practitioners are used to model the national mortality and then to project the insured ones. One example of how such projections are produced is by relational models : a technique that aims at linking together both mortality rates and their future evolution, even if socio-economic variables cause the insurance portfolio specific mortality and the national mortality to differ.

In addition, at the portfolio level, whilst the movements in the group of policyholders are known and can be taken into consideration, it proves difficult to account for censored and truncated data when assessing national mortality. One may think for instance of migration. There are two different factors that can be overlooked when considering migratory impacts. First of all, people leaving the synthetic cohort are often censored, and thus still taken into account when estimating the period life tables. Secondly, incoming immigrants can also affect and significantly change the local national mortality data of the destination country. Far more relevant details regarding deaths can be gained from within an insured portfolio (e.g. the cause of death). Although national data by cause of death does exist, it generally lacks consistency, and is often useless, or of no specific interest in deriving mortality by cause (for example : in the United States, 11% of deaths are caused by more than four diseases).

1. The available mortality data is based upon statistics coming from various national institutes (Including but not limited to : Bureau of Census in the US, CMI in the UK, INSEE in France). For most developed countries, this data is also available through the Human Mortality Database.

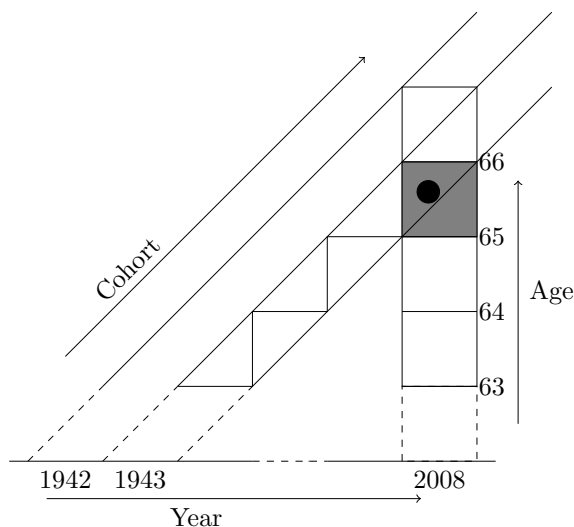


FIGURE 4.1 – Lexis diagram : Age-year-cohort diagram.

4.1.2 Observed heterogeneity, inter-age dependence and basis risk

It is expected that any given population will display some degree of heterogeneous mortality. Observed heterogeneity often arises due to a number of observable factors, which include age, gender, occupation and physiological factors. As far as longevity risk is concerned, policyholders that are of higher socio-economic status (assessed by occupation, income or education) tend to experience lower levels of mortality. However, significant differences also exist within the same socio-economic levels according to gender. Females tend to out-live males and have lower mortality rates at all ages. In addition, some heterogeneity arises due to features of the living environment, such as : climate, pollution, nutritional standards, population density and sanitation (see Section 4.2.2 for a more detailed discussion).

When considering insured portfolios, insurers tend to impose selective criteria that limits contractual access to those considered to possess no explosive risk (by level of health and medical history). This induces differences in mortality profiles within different portfolios. For example, term and endowment assurance portfolios are experiencing higher mortality rates than those related to annuity portfolios and pensions. Also, the risks underlying those two types of contracts are different in nature. The resulting risk for the prior portfolios is a mortality risk, whilst the latter types have a longevity risk. By virtue of their opposing nature, the possibility remains for a hedge that implies both types of risk². However, as highlighted by [17], this natural hedge is only partial. The hedging is only partially attain-

2. when mortality rates decrease, life annuity costs increase and death benefit costs decrease and vice versa.

nable due to the nature of both risks and the different age ranges implied in the two specific contracts. In particular, the interdependence among ages, and even the inter-temporal correlation are dynamically significant and need to be understood.

In [38], both the issue of inter-age dependency according to the mortality data and the inter-temporal correlation are considered. In this study, the population data for both males and females exhibits a clear dependence upon mortality rates among ages. This inter-age dependency is crucial when dealing with, for example, the natural hedge between mortality and longevity. More precisely, it gives a clear view of how some changes in mortality for a specific age-range will affect another by providing measurements of the associated liabilities altered by potential diversification due to inter-age correlation. Thus, it is a characteristic of considerable interest and importance when aggregating the benefits within a book of contracts.

4.2 Modelling longevity risk

In the previous section a detailed description of longevity risk and the available data has been presented. The next section will discuss the implications and issues of modelling, outlining some key methods that can be used to take into account relevant characteristics.

4.2.1 Some standard models

In recent years a variety of mortality models have been introduced, including the well known ([37]), widely used by insurance practitioners. The Renshaw-Haberman model ([30, 30]) was one of the first to incorporate a cohort effect parameter to characterise the observed variations in mortality among individuals from differing cohorts. For a detailed survey on the classical mortality models, [41] is referred to. In recent contemporary studies, many authors have introduced stochastic models to capture the cohort effect (see e.g. [11, 13]); this subsection, briefly presents some of them.

The Lee-Carter model The Lee Carter model describes the central mortality rate $m_t(x)$ or the force of mortality, $\mu_{x,t}$ at age x and time t by 3 series of parameters namely α_x , β_x and κ_t , as follows :

$$\log \mu_{x,t} = \alpha_x + \beta_x \cdot \kappa_t + \varepsilon_{x,t}, \quad \varepsilon_{x,t} \sim \mathcal{N}(0, \sigma) \text{ and independent,}$$

where α_x gives the average level of mortality at each age over time, the time varying component κ_t is the general speed of mortality improvement over time, and β_x is an age-specific component that characterises the sensitivity to κ_t at different ages. To enforce the uniqueness of the parameters, some constraints are imposed :

$$\sum_x \beta_x = 1 \quad \text{and} \quad \sum_t \kappa_t = 0.$$

The β_x also describes (on a logarithmic scale) the deviance of the mortality from the mean behaviour κ_t . The error term $\varepsilon_{x,t}$ captures the remaining variations.

Some relevant extensions of this approach have also been proposed, incorporating, for example, the generation or cohort effect as in [30] or [31] describing the influence of the year of birth. Meanwhile, given the high variability of the force of mortality at higher rather than at younger ages (due to a much smaller absolute number of deaths at older ages), the error term cannot be regarded as being correctly described by a white noise. [10] has proposed an extension of the general framework by assuming the number of deaths to be a Poisson distribution. To calibrate the various parameters, standard likelihood methods can be applied, and a Poisson distribution can be considered for the number of deaths at each age over time. The estimated parameters α , β and κ for the English and Welsh population are presented on Figure 4.2 and Figure 4.3. It is important to note that estimated values for β_x are higher at younger ages, meaning that at those ages, the mortality improvements occur faster and deviate considerably from the evolutionary mean.

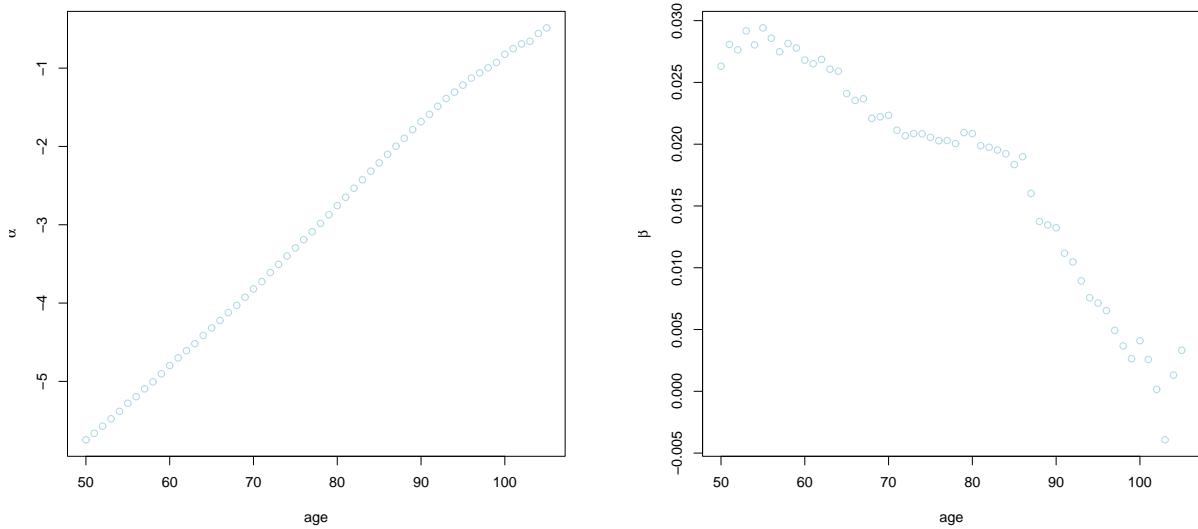


FIGURE 4.2 – Parameters estimates for the England and Wales mortality table

In addition to the [37] model, similar models have been introduced to forecast mortality. For instance, the [11] two-factor model includes two time dependent components rather than a unique component as seen in the original Lee & Carter model. This model could be thought of as a compromise between the generalised regression approach and the Lee & Carter method. In addition, this approach allows components relevant to the cohort effect to be incorporated (for further details see [13]).

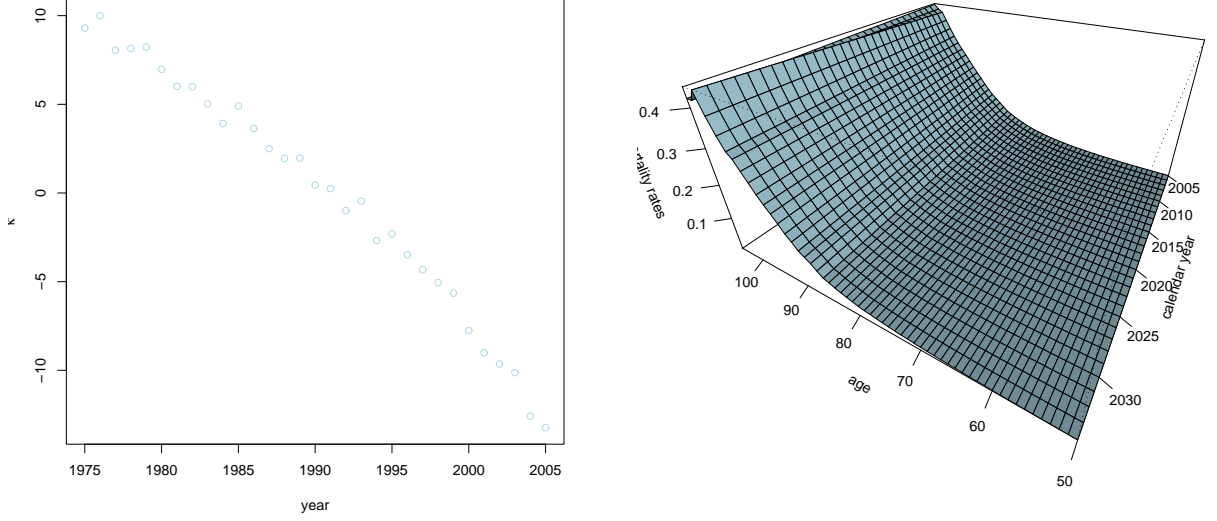


FIGURE 4.3 – Parameters estimates for the England and Wales mortality table

The P-Spline model The P-spline model is widely used to model UK mortality rates. The mortality rates are incorporated by the use of penalised splines (P-splines) that derive future mortality patterns. This approach is used by [18] to smooth the mortality rates and extract "shocks" as considered by [19], which can be used to derive stress-test based scenarios. Generally, in the P-spline model, the central mortality rate $m_t(x)$ at age x and calendar year t satisfies :

$$\log m_t(x) = \sum_{i,j} \theta^{i,j} B_t^{i,j}(x),$$

where $B^{i,j}$ are the basis cubic functions used to fit the historical curve, and $\theta^{i,j}$ are the parameters to be estimated. A difference between the P-spline approach and the basis cubic spline approach is found when introducing penalties on parameters $\theta^{i,j}$ to adjust the log-likelihood function. To predict mortality, the parameters $\theta^{i,j}$ have to be extrapolated using the given penalty.

The CDB model Cairns, Dowd and Blake (CDB) have introduced a general class of flexible mortality model that takes into account different purposes and underlying shapes of mortality structures. The general form of the annual death probability $q_t(x)$ at age x and calendar year t is given by :

$$\text{logit} q_t(x) = \kappa_t^1 \beta_x^1 \gamma_{t-x}^1 + \cdots + \kappa_t^n \beta_x^n \gamma_{t-x}^n, \quad (4.1)$$

where $\text{logit}(q) = \ln\left(\frac{q}{1-q}\right)$. As can be seen, there are 3 types of parameters : starting with those specific to age β^i and calendar year κ^i , and finally, the cohort effect parameters γ^i . It

should be noted that the Lee-Carter model is a simple case of this model. The authors CBD also investigate the best criterion to decide upon a particular model (i.e. the parameters to keep or to remove). Thus, they highlight the need for a tractable and data consistent model ; forming statistical gauges to rank models, determining their suitability for mortality forecasting. A particular example of a model derived from the general form (4.1) is the model below, which features both the cohort effect and the age-period effect, e.g. [13] :

$$\text{logit}\mu_{x,t} = \kappa_t^1 + \kappa_t^2(x - \bar{x}) + \kappa_t^3 \left((x - \bar{x})^2 - \sigma_x^2 \right) + \gamma_{t-x},$$

where

$$\bar{x} = \frac{\sum_{x=x_0}^{x_n} x}{x_n - x_0 + 1}$$

is the mean age of historical mortality rates to be fitted (x_0 to x_n) ; σ_x is the standard deviation of ages, with

$$\sigma_x^2 = \frac{\sum_{x=x_0}^{x_n} (x - \bar{x})^2}{x_n - x_0 + 1};$$

the parameters κ_t^1 , κ_t^2 and κ_t^3 correspond respectively to general mortality improvements over time, the specific improvement for specific ages (taking into account the fact that mortality for high ages improves slower than for younger) ; and finally the age-period related coefficient $\left((x - \bar{x})^2 - \sigma_x^2 \right)$, which corresponds to the age-effect component. Similarly, γ_{t-x} represents the cohort-effect component.

4.2.2 Observed heterogeneity

The models introduced above are generally applied to large observed datasets, and thus, are most suited to modelling mortality rates for an underlying national population. However, as discussed above, insurers have a keen interest in the mortality rates experienced in their own portfolios ; any discrepancy between both sets of mortality figures represents a significant risk for the insurer : this is referred to as basis risk.

Analysis of the basis risk

[39] propose a model that links the insured specific mortality and national population mortality using an econometric approach that captures the long-run relationship of the behaviour of both mortality dynamics. Rather than emphasising the correlation between both mortality dynamics, the model focuses on the long-term behaviour. This suggests that both time-series cannot diverge for any significant amount of time, without eventually returning to a mean distance. This kind of model, based on a co-integration relationship, is only effective if a long enough period of observable data can be used to detect any long-term relationships ; thus, is most appropriate for insured portfolios with a large observation history. Relational models (see [21]) offer an alternative when the insured specific data is

observed for a limited period.

For a complete analysis of the basis risk, it is important to take into account individual characteristics, such as relative socio-professional status, individual income, education, matrimonial status and other factors. In fact, according to some demographers who took part in the workshop "Demographic perspectives : How to measure and deal with uncertainty"³, any slump in European mortality would principally result from developments in medical research, and at the socio-economic level (including increased access to medical care, education and other quality of life improving factors). In addition, a recent study published by researchers from IRDES⁴ and INED⁵, describes a strong relationship between individual socio-economic level and life expectancy (see [35]). For example, the difference in life expectancy between two French males at 35 years who occupy opposite bounds of the socio-professional spectrum - one a manual worker and the other an office executive - is 6 years (see Cambois and al 2008). However, as emphasised in Subsection 4.1.1, a model that describes mortality by causes such as diseases is limited by a lack of both reliability and objective mortality data : and is ultimately unable to accurately and consistently identify the cause of death. The classical mathematical models for mortality, such as the models presented in Subsection 4.2.1, consider mortality rates to be a stochastic process, dependent exclusively upon age, calendar year, and date of birth. To analyse basis risk, one needs to extend these models. One new approach is microscopic modelling ; a method that accurately describes characteristics of a population at the scale of the individual ([5]). Inspired by the Cairns-Dowd-Black methodology presented in Subsection 4.2.1, a model can begin to take shape that has an age dependent mortality rate, but is additionally inclusive of various finer characteristics of both individuals and the associated environment. The construction of such a model requires a study identifying which individual characteristics (other than age) could explain mortality, with the characteristics eventually incorporated into a stochastic mortality model. This approach reduces the variance of the mortality rate by taking into consideration specific information about the studied population. From a financial point of view, this model gives accurate information on portfolio basis risk.

Example of heterogeneity based on the matrimonial status

In order to illustrate the influence of matrimonial status on mortality, it is useful to consider the following study of the application of longitudinal analysis upon multi-census data from 1968 to 2005 provided by the French institute INSEE⁶. The data corresponds to a sample of about 1% of the French population. By applying a longitudinal analysis to this data, one can derive mortality rates that are dependent upon age and additional features

3. see <http://www.dauphine.fr/fileadmin/mediatheque/chaire/chairetransition/pdf/Rapport2009.pdf>

4. Institut de Recherche et Documentation en Économie de la Santé

5. Institut National d'Études Démographiques

6. Institut National de la Statistique et des Études Économiques

such as socio-professional categories and education. This method can also be used to derive matrimonial status, which is the feature that will be focused on.

To reach a greater level of detail, the evolutionary development of the logistic transform for mortality by age and by matrimonial status will be considered. There is a considerable difference in mortality rates between married and unmarried (single, divorced, widow) individuals. Marriage is known to have protective qualities, which is clearly illustrated by the surge of mortality rate for unmarried people which can be seen on the left hand side of Figure 4.4.

It is also possible to predict a mean evolution of these curves to reflect points in the future; as such, the second graph of Figure 4.4 represents modelled data that expresses the expected evolution of mortality in 2017, giving an aggregate mortality profile of the whole population. Furthermore, the mean evolution is comparable to the pattern of mortality by using the modelling processes described in Subsection 4.2.1.

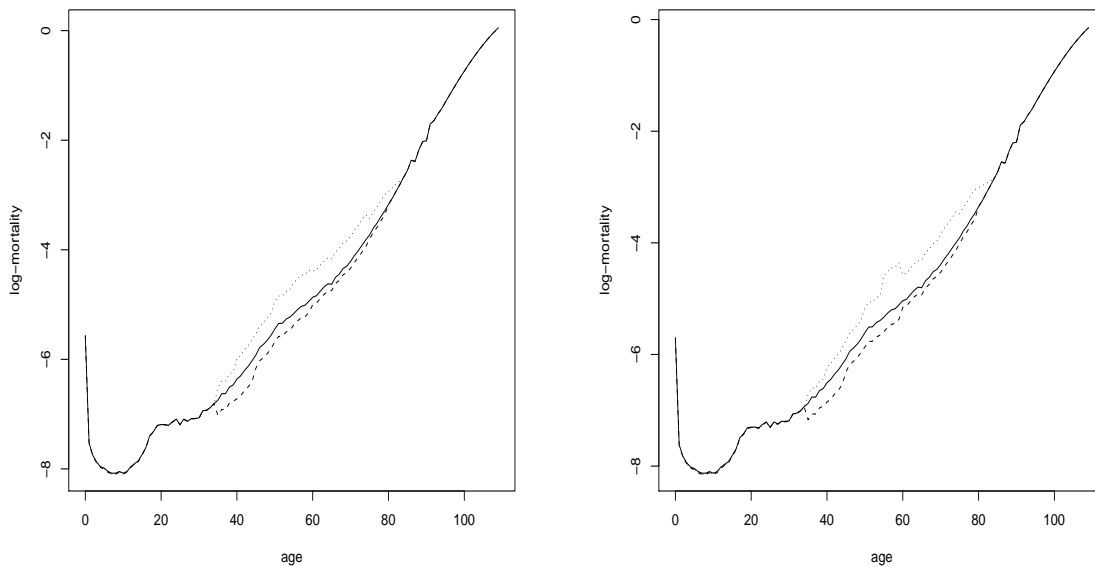


FIGURE 4.4 – Logit of mortality rate for French males in 2007 (left) and in 2017 (right) with differing matrimonial status : National (solid), unmarried (dotted) and married (dashed)

Drawing upon recent probabilistic research ([24], [43] and [44]) and considering models that reflect mortality rates, a population dynamic modelling process is proposed in [5]. This model takes into account individual mortality rates and provides projections of a population structure for a forthcoming year. A mean scenario of evolution can be deduced and analysed from these simulations; but extreme scenarios and their probability of occurrence have also to be taken into consideration. This model can be used for the study of basis risk, allowing for a finer assessment of mortality and giving the heterogeneity factors affecting

each individual.

A specific example of the use of such a framework can be seen in annuity products and pensions. When considering a portfolio of life annuities for 10000 60 year old French males, the marital status of the annuitants can be distinguished. Given that mortality correlates with matrimonial status depending on age and the cohort, the value of the portfolio will change according to the individual characteristics. For example, considering such a portfolio, it is possible to simulate the central scenario for the cash flows the annuity provider is liable to pay according to differing matrimonial status. Figure 4.5 summarises the cash flow structure according to individual matrimonial status observed in 2007. For more details on the aforementioned modelling, and for the complete study detailing the influence of additional features, such as socio-professional categories and education : refer to [5].

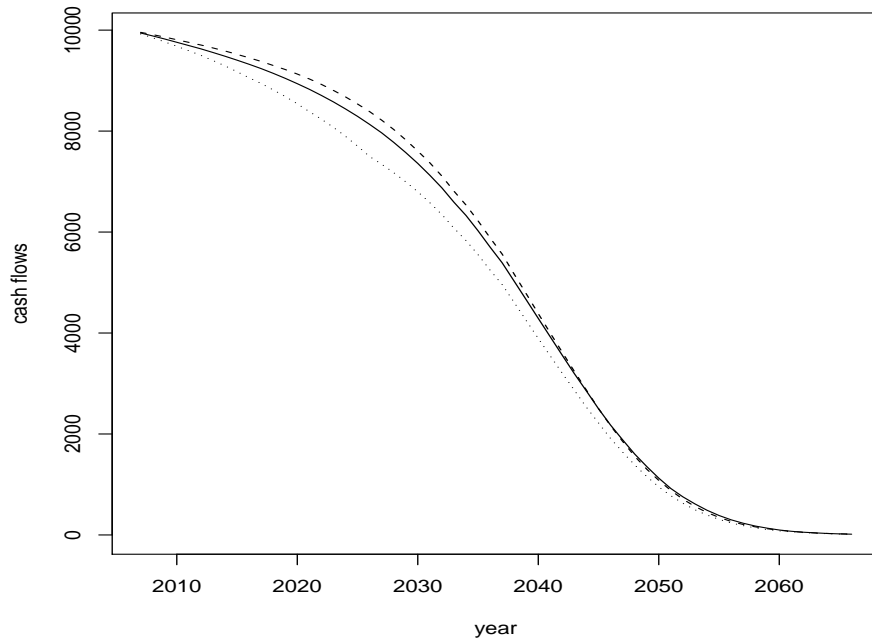


FIGURE 4.5 – Cash flows of a life annuities portfolio for 60 year old French males considered by central scenario (median scenario) for different matrimonial status : National (solid), unmarried (dotted) and married (dashed).

4.2.3 Modelling inter-age dependence and inter-temporal correlations

As mentioned in the previous section, an important practical issue that arises when dealing with life insurance portfolios is the interaction among the different ages that constitute the portfolio. In practice this problem has to be taken into account, and to do so involves a multi-dimensional analysis. In [38], the model proposed is a multivariate autoregressive

ARIMA model, which accounts for both inter-age and inter-period correlation. This model provides a more sophisticated approach for assessing, for example, the effect of diversification within a multi-age life insurance portfolio. However, the high number of parameters that requires estimation in such a framework can potentially lead to tractability issues. [7] proposes a more tractable framework with aggregate mortality for a class of ages. Inter-age correlations turn out to be significantly positive, and less than 1. As a consequence, they have to be quantified carefully to compute the prices of longevity transfer contracts, and not least, to determine internally modelled solvency capital requirements required by the new regulatory guidelines.

4.3 New regulations, economic and sociologic impacts of longevity risk

The establishment of the new regulatory norms, namely Solvency II, stands to ensure the establishment of more accurate standards, unifying and homogenising solvency capital computation and risk assessment practices. These regulations are particularly significant to actuarial practices in life insurance, primarily because their practices are based on a deterministic view of risk. Although such practices are indeed prudent and ensure the solvency of the insurer, they exclude any unexpected risk deviation. In fact, the amount of provisions, and the value of products themselves are, in most cases, obtained via deterministic computational methods. Thus, the calculation of provisions is reduced to a net-present-value of future cash flows discounted with risk-free rates. The new standards highlight the pressing need for the market price of risk to be incorporated into both the calculation of provisions, and the evaluation of products to create the desired result of "market consistent" values. For this purpose, regulators differentiate between two kinds of risk : hedgeable risk and non-hedgeable risk. The latter is widely discussed and treated independently of any market. For hedgeable risks, however ; the hedging strategy is used to evaluate the underlying liabilities. The following section will briefly outline the differing calculations - on technical provisions, and capital requirements - used for longevity and associated risk factors. In addition, some pitfalls relating to the use of the Solvency II standard formula will be discussed.

4.3.1 Technical provisions

The technical provisions stand for the anticipated liabilities, and as such, are reported on the liability side of the insurer's balance sheet. The Solvency II directive proposals and, more precisely, the quantitative impact studies (QIS) (see [15] for more details), are proposing standards that will unify practices of provisions calculation and product valuation. In particular, technical provisions will have to be calculated by taking into account the available market information. In other words, the provisions should be market-consistent. They are computed as the best estimate of future liabilities plus a risk margin. The estimate

of future liabilities is based on realistic assumptions about the future evolution of various risk factors. More precisely, the risk factors are first estimated and then the relevant future patterns are derived under a set of prudential assumptions. In this case, the best estimated value of a liability is simply the mean over all future scenarios.

Both in practice and for longevity linked contracts, the best estimate assumptions could be derived from internal models, or based upon relevant models that allow for the identification of future mortality patterns. The models presented in the previous section are suitable for this purpose.

The fact that the best estimate does not replicate the actual value of a liability, imposes constraints upon insurers : necessitating the holding of an excess of capital to cover the mismatch between the best estimate and the actual cash flows of the liability. Such capital is referred to as the Solvency Capital Requirement (SCR).

4.3.2 Solvency Capital Requirements

As has been mentioned earlier, any insurer must constitute some reserves, including the SCR, to ensure its solvency. According to QIS4, the SCR will require a firm's solvency standing to be equivalent to a BBB rated firm. In other words, *"equivalent to the firm to hold a sufficient capital buffer to withstand a 1 in 200 year event (the otherwise termed 99.5% level)"*.

The calculation of the SCR will rely on either an internal model that captures the firms risk profile, or standardised formula proposed by the fifth quantitative impact study (QIS5) ; and as such will determine risk profiles by the use of a variety of 'modules'.

Standard formula approach and its drawbacks

In the standard formula, the capital calculation is computed separately for each module and risk factor, and then aggregated.

Firstly, there is the module-based framework that proposes pre-defined scenarios to compute solvency capitals. Concerning the longevity risk, capital requirements must be added to the best estimate technical provisions to both withstand unexpected deviations of the mortality trends, and allow the insurer to meet any obligations in adverse scenarii. For this purpose, insurers should use a scenario-based method that involves permanent changes in mortality improvement factors $\frac{q_{t+1}(x) - q_t(x)}{q_t(x)}$. For example, the proposed mortality risk scenario includes a permanent 10% increase in mortality improvement from the baseline forecast. Similarly, for contracts that provide benefits over the whole life of the policyholder (i.e. longevity risk), the scenario suggests that an additional permanent 20% decrease for mortality improvement factors should be set each year.

Finally, the entire solvency capital is aggregated, given by the equation :

$$\text{SCR}_{\text{global}} = \sqrt{\sum_{i>0} \sum_{j>0} \theta_{i,j} \text{SCR}_i \text{SCR}_j},$$

where the set of correlation parameters $\Theta = (\theta_{i,j})_{i>0, j>0}$ is prescribed by the regulator. The regulators seem to recognise the existence of a 'natural' hedge between the components of mortality and longevity risk. This natural hedge can be translated to be the correlation parameter between the longevity and mortality risk modules, which is assumed to be negative and equal to -25% . It must be noted, that this is the only negative correlation parameter in all QIS5 correlation matrices. But how do longevity and mortality risks correlate in reality? Longevity appears as a trend risk, whereas mortality is variability risk. Is there orthogonality between mortality and longevity? - In other words, can we "buy" mortality risk in order to hedge longevity risk? Oscillations around the average trend are also important : the size of the oscillations cannot be neglected, and in any case they can lead to over-reactions by insurance managers, regulators, policyholders and governments. Obviously, even if a certain mutualisation between mortality and longevity risks exists, it is very difficult to obtain a significant risk reduction between the two, on account of their differing natures. Indeed, the replication of life annuities with death insurance contracts could never be perfect as it concerns different individuals; thus, the hedge is often inefficient due to the variability related to the death of insured individuals whose death benefits will be high. Moreover, the impacts of a pandemic or a catastrophe on mortality is considerably different from their impacts on longevity. In fact, an abnormally high death rate at a given date has a qualified influence on the longevity trend, as it was with the 1918 flu pandemic (see Figure 4.6).

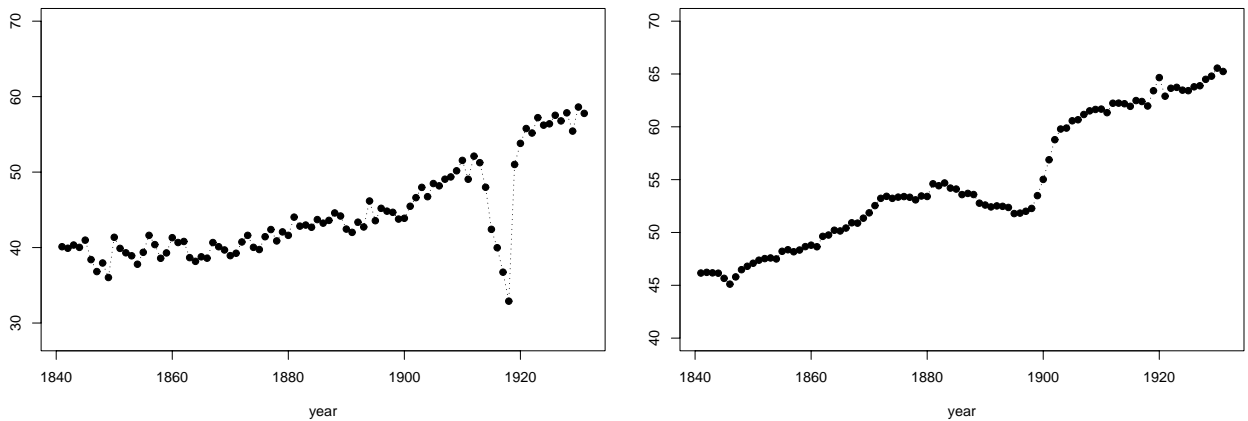


FIGURE 4.6 – Periodic life expectancy (left) and generational life expectancy (right) at birth in the UK.

What about dependence between life and market risks ? Long-term horizons have financial consequences : if longevity risk is transferred, the interest rate and counterparty risk often become predominant. Financial risks are grouped into the market risk module ; whereas, the longevity, mortality and disability risks are grouped into the life risk module. At a higher level, market-risk and life-risk are aggregated in the same manner thanks to a unique correlation parameter that excludes data on the proportion of longevity and mortality risk companies are exposed to. This is a major deficit of the two-level aggregation approach as described in [22].

The correlation parameter between longevity risk and disability risk, fixed at the level -25% , is relatively low in QIS5. Is it realistic in practice ? The loss of autonomy and increased dependency that often affects elderly people, are major demographic issues. Taking France as an example, important approaches have been introduced in order to help so called "dependent people". The question of long-term care insurance has a strong economic dimension, with a year of dependence costing four times more than a typical year of retirement. Consequently, a worldwide public awareness campaign about the phenomenon and its consequences is being run to promote preventative action. Moreover, levels of dependence accurately reflect individual health capital. Statistics from the INSEE reveal that when an individual enters a state of dependence, their life expectancy drops by 4 years, almost entirely irrespective of age. The INSEE statistics were obtained by factoring in all states of dependency, from the chronic dependencies such as those linked to Alzheimer's disease, to the less severe forms, that often account for several years. Thus, individual longevity is strongly correlated to dependence level. Obtaining an idea of the global correlation between the two risks at portfolio level is difficult and requires modelling similar to that presented in [5]. Finally, it is important to note that the meaning and definition of "dependence" (or "disability") differs from country to country and across regulatory panels, making the correlation between longevity and dependence difficult to define.

Thus, insurers have an incentive to develop an internal model, or a partial internal risk model, to better account for the characteristics and interactions of biometric and financial risks.

Internal risk models

The alternative to the standard formula for calculating solvency capital requirements is the use of an internal model (or one or more partial models). In such cases, the internal model should capture the risk profile of the insurer by identifying the various risks it faces. Therefore, the internal model should incorporate the identification, measurement and modelling of the insurer's key risks. As for internal models, the Solvency II guidelines take the position that in circumstances where an insurer prefers developing its own framework to assess the incurred risks, the Value at Risk must be used to compute the required capital. The methodology considered here is highly contrastable with the one already in use in the

banking industry.

The Value at Risk measure that has been recently introduced within the insurance industry, is based on yearly available data. This highlights the main difference between the banking and insurance industries. In banking, access is assured to high frequency data that permits computation of daily risk measures; whereas, in insurance the Value at Risk is computed over the whole year, and so can assess solvency. The required capital for the year SCR_i , insuring the solvency during this given period, is drawn from the Value-at-Risk at level $\alpha = 0.5\%$ as

$$SCR_i = VaR_\alpha(M_i) - \mathbb{E}(M_i),$$

where M_i is the liabilities used to compute the associated solvency capitals.

4.3.3 Risk margin

Finally, the required capital has to yield a return (it is not necessarily fixed, and it depends on the internal capital targets) each year, as it is sold to shareholders who expect some return on their investment. The total margin needed to satisfy the return expected by the shareholders, is called the risk margin. Another method, called the cost of capital approach, is used to determine the market price of risk, and possible transfer values when added to the best estimate of liabilities. Regulators have highlighted the effectiveness of risk mitigation such as reinsurance and derivatives to release capital, which is especially important considering the new additional solvency requirements increase the need for capital. Therefore, capital markets are an attractive means by which to transfer the longevity risk, as the traditional reinsurance approach to risk transfer has a limited capacity to fund and absorb that risk.

4.3.4 Impact of longevity risk on the economy

To think beyond the Solvency II framework, a risk management analysis that incorporates dynamic correlations between both life and financial risks should also take into account the impacts of longevity risk upon the economy. As has been noted, many entities are concerned by longevity risk, and will find it necessary to hedge the long-term risk. For example, governments are now facing a retirement challenge coupled with associated longevity risk. The article [1] highlights many of the impacts caused by longevity improvements for people aged eighty or older on GDP and political decision-making at a national level. Consequently, population ageing has macroeconomic consequences; some countries are even considering it as a factor of economic slackening.

However, this issue can be mitigated: when life expectancy increases, consumption also increases. The ageing of a population does not necessarily correspond to an economic 'ageing' and associated decline, but could, on the contrary, inspire an economy of ageing. Innovations come in many fields, including but not limited to: medicine, home automation, urban

planning and transport. In many developed countries, urban redevelopment incorporates the facilitation of many access types, including free circulation for the elderly.

In 2005, the French Academy of Pharmacy published a report "Personnes Âgées et Médicaments" (see [23]) that revealed an increase in medicine consumption by both senior citizens and the whole medical economy of ageing. Indeed, with an ageing population and growing demand, medicine consumption increases at a high rate. Moreover, the innovative pharmaceutical companies are considering the development of new medicines specifically designed for the elderly.

4.4 Transferring longevity risk

To recap, a steady increase in life expectancy in Europe and North America has been observed since 1960s, which has and continues to represent a significant and evolving risk for both pension funds and life insurers. Various risk mitigation techniques have been advocated to better manage this risk : reinsurance and capital market solutions in particular, have received growing interest. In February 2010, with the aim of promoting and developing a transparent and liquid market for longevity risk transfer solutions : AXA, Deutsche Bank, J.P. Morgan, Legal and General, Pension Corporation, Prudential PLC, RBS and Swiss Re established the Life and Longevity Markets Association (LLMA). The LLMA is a non-profit venture supporting the development of consistent standards, methodologies and benchmarks.

4.4.1 Convergence between insurance and capital markets

Even if no Insurance-Linked Securitisation (ILS) of longevity-risk has so far been implemented, the development of the market for other insurance risks has been experiencing continuous growth for several years. This growth that has been mainly driven by changes in the regulatory environment and the distinct need for additional capital in the insurance industry. Today, longevity risk securitisation lies at the heart of many discussions, and is widely seen as a viable potential.

The convergence of the insurance industry with the capital markets has become increasingly important over recent years. This convergence has taken many forms, and with mixed success. In academic discourses, the first theoretical discussion presenting the idea that capital markets could be effectively used to transfer insurance risk was presented in a paper by [27]. The authors considered the feasibility of an organised market, questioning how such a market could complement the reinsurance industry, discussing primarily catastrophic risk management. In practice, while some attempts have been made to develop an insurance futures and options market, so far the results have caused some disappointment. Despite this, the ILS market has grown rapidly over the last 15 years. Many motivations exist for using ILS instruments, including : risk transfer, capital strain relief, acceleration of

profits, speed of settlement, and duration. Different motivations, require different solutions and structures, as the variety of instruments on the ILS market illustrates.

While the non-life section of the ILS market is the most visible, famously trading the highly successful cat-bonds, the life section of the ILS market is the bigger in terms of transaction volume. Today's situation is varied, and there are huge contrasts between the non-life and the life ILS markets ; with success, failure and future developments resting with the impacts of the financial crisis. Whilst in the non-life-sector, the credit crisis has had only a limited noticeable impact, partly due to product structuring, a dedicated investor base and a disciplined market modelling and structuring practices, the life-sector has been greatly affected by the recent crisis. This is mostly due to the structuring of deals and the nature of the underlying risks : with more than half of the transactions being wrapped, or containing embedded investment risks. Therefore, principles governing the constitution and management of the collateral account, as well as the assessment of the counterparty risk are central to current debates aimed at developing a sustainable and robust market.

4.4.2 Recent developments in the transfer of longevity risk

Coming back to longevity risk ; a number of important developments have been observed over the past 3 years, including increased attention from US and UK pension and life insurance companies, and an estimated underlying public and private exposure of over 20 trillion USD. Even if a large quantity of private equity transactions have been completed, very few overt capital market transactions - mainly taking a derivative form (swaps) - have been achieved.

Despite this limited activity, using the capital markets to transfer part of the longevity-risk is complementary to traditional reinsurance solutions, and would thus seem to be a natural move. To date, almost all longevity capacity has been provided by the insurance and reinsurance markets. Whilst this capacity has facilitated demand to date, exposure to longevity risk for UK pension funds alone is estimated to exceed 2 trillion GBP. It is therefore clear that insufficient capacity exists in traditional markets to absorb any substantial portion of the risk, and only capital markets are a potential capacity provider.

On face value, longevity meets all the basic requirements of a successful market innovation. However, there are some important questions to consider. To create liquidity and attract investors, annuity transfers need to move from an insurance format to a capital markets format.

As a consequence, one of the main obstacles for the development of capital market solutions would appear to be the one-way exposure experienced by investors, since there are almost no natural buyers of longevity-risk. Inevitably, this could cause problems for demand. Nevertheless, provided it is priced with the right risk-premium, there is potential for longevity-risk to be formed as new asset class, which could interest hedge funds and specialised ILS investors.

Another consideration is that basis-risk could prevent a longevity market from operating successfully. Indeed, the full population mortality indices have basis risk with respect to the liabilities of individual pension funds and insurers. Age and gender are the main sources of basis risk, but also regional and socio-economic basis risk could be significant. Therefore, the use of standardised instruments based upon a longevity index, with the aim of hedging a particular exposure, would result in leaving the pension fund or life insurer with a remaining risk often challenging to understand (and hence, to manage). An important challenge lies in developing transparency and liquidity by ensuring standardisation, but without neglecting the hedging purposes of the instruments.

Recently, many different initiatives have been undertaken in the market, which aim at increasing the transparency concerning longevity-risk and contributing to the development of longevity-risk transfer mechanisms.

4.4.3 Various longevity indices

Among the different initiatives to improve visibility, transparency and understanding of longevity-risk, various indices have been created. A longevity-index needs to be based on national data (available and credible) to have some transparency, but must also be flexible enough to reduce the basis-risk for the original longevity-risk bearer. National statistical institutes are in a position to build up annual indices based on national data, which could incorporate projected mortality rates or life expectancies (for gender, age, socio-economic class and so on). Potentially, such a move could limit basis-risk and help insurance companies form a weighted average index pertinent to their specific exposure.

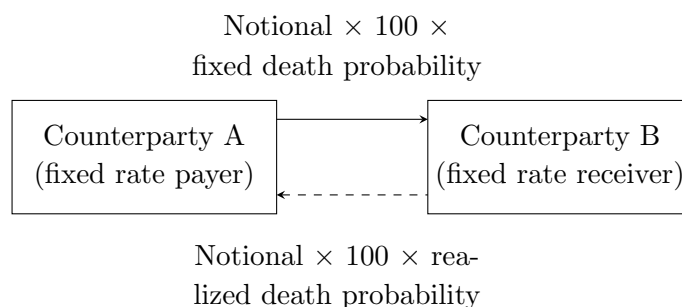
Today, the existing indices⁷ are :

- Credit Suisse Longevity Index, launched in December 2005, is based upon national statistics for the US population, incorporating some gender and age specific sub-indices.
- JP Morgan Index with LifeMetrics, launched in March 2007. This index covers the US, England, Wales and the Netherlands, by national population data. The methodology and future longevity modelling are fully disclosed and open, (based upon a software platform that includes the various stochastic mortality models).
- Xpect Data, launched in March 2008 by Deutsche Borse. This index initially delivered monthly data on life expectancy in Germany, but has now been extended to include the Netherlands.

7. The Goldman Sachs Mortality Index, launched in December 2007, was based on a sample of insured over 65s in the U.S. It targeted the life settlement market, but was discontinued in December 2009, following the global financial crisis.

4.4.4 q-forwards

JP Morgan has been particularly active in trying to establish a benchmark for the longevity market. Not only have they developed the LifeMetrics longevity risk platform, but have developed standardised longevity instruments called "q-forwards". These contracts are based upon an index that draws upon either the death probability or survival rate as quoted in LifeMetrics. Naturally, survivor swaps are the intuitive hedging instruments for pension funds and insurers. However, the importance assigned to the starting date of the contract (owing to the survival rate being path-dependent), may reduce the fungibility of the different contracts in relation to the same cohort and time in the future. Therefore, mortality swaps are also instruments likely to be applicable. The mechanisms of a q-forward can be summarised as follows :



Notional Amount	GBP 50,000,000
Trade Date	31 Dec 2006
Effective Date	31 Dec 2006
Maturity Date	31 Dec 2016
Reference Year	2015
Fixed Rate	1.2000%
Fixed Amount Payer	JPMorgan
Fixed Amount	$\text{Notional Amount} \times \text{Fixed Rate} \times 100$
Reference Rate	LifeMetrics graduated initial death probability for 65-year-old males in the reference year for England & Wales national population
Floating Amount Payer	WYZ Pension
Floating Amount	$\text{Notional Amount} \times \text{Reference Rate} \times 100$
Settlement	Net settlement = Fixed amount - Floating amount

FIGURE 4.7 – An example of a q-forward contract

The mechanisms of the q-forwards can be understood simply : a pension-fund hedging its longevity-risk will expect to be paid by the counterpart of the forward if the mortality falls by more than expected. So typically, a pension fund is a q-forward seller, while an investor is a q-forward buyer.

4.4.5 Longevity swap transactions and basis risk

Over the last 3 years, a number of longevity swap transactions have taken place. The transactions have occurred in private ; therefore, their pricing remains confidential and subject to the negotiation between the parties involved in the deal. Some swaps were contracted between a life insurance company and a reinsurer as a particular reinsurance agreement ; whilst, others have involved counterparts outside the insurance industry. Most of these transactions have a long-term maturity and incorporate an important counterparty risk, which is difficult to assess given the long term commitment. As a consequence, the legal discussions around these agreements make them a particularly drawn out contract to finalise. Longevity swaps mainly take two forms, depending on whether they are index-based or customised. Two differing longevity swaps arranged by JP Morgan in 2008 are discussed below :

A customised swap transaction In July 2008, JP Morgan executed a customised 40 year longevity swap with a UK life insurer for a notional amount of GBP 500 million. The life insurer agreed to pay fixed payments, and to receive floating payments, replicating the actual benefit payments made on a closed portfolio of retirement policies. The swap is before all a hedging instrument of cash flows for the life insurer, with no basis risk. At the same time, JP Morgan entered into smaller swaps with several investors that had agreed to take the longevity risk at the end. In this type of indemnity based transaction, the investors have access to the appropriate information to enable them to assess the risk of the underlying portfolios. The back-to-back swap structure of this transaction means that JP Morgan has no residual longevity exposure. The longevity risk is transferred from the insurer to the investors, in return for a risk premium. The counterparty risk for such a swap is important given the long term maturity of the transaction, and the number of agents involved.

A standardised transaction : Lucida In January 2008, JP Morgan executed a 10 year standardised longevity swap with the pension insurer Lucida for a notional sum of GBP 100 million ; with an underlying risk determined by the LifeMetrics index for England and Wales. This swap structure enabled a value-hedge for Lucida, who have in this case agreed to keep the basis risk. For more details on both transactions, and longevity swaps, see [2].

As previously outlined, longevity-risk is highly particular. The market for longevity-risk is also very distinct, being strongly unbalanced in terms of exposures and needs. This

makes the question of the pricing of risk transfer solutions markedly important. Yet, even before this, the problem of designing suitable, efficient and attractive structures for both risk bearers and risk takers is absolutely essential to tackle. This is underlined by the failure of the EIB-BNP Paribas longevity bond⁸ in 2005. The recent financial crisis has also emphasised the importance of assessing counterparty-risk, and properly managing the collateral accounts which help to secure transactions. These questions are even more critical when considering longevity-risk : firstly, due to the long-term maturity of the transactions ; but secondly, the social, political and ethical nature of the risks involved.

4.5 Modelling issues for pricing

Designing longevity-based securities brings together various modelling issues besides form the challenges of pure longevity risk modelling. An initial issue, is that the pricing of any longevity "derivative" is not straightforward, as it is dependent upon both estimations of uncertain future mortality-trends, and other levels of uncertainty. This risk induces a mortality-risk premium that should be priced by the market. However, because of the absence of any liquid traded longevity-security, it is today impossible to rely on market data for pricing purposes. Long-term interest rates should play a key role in the valuation of such derivatives with long maturities (up to 50 years), which is likely to create new modelling challenges.

4.5.1 Pricing methodologies

In this section, methodologies for pricing longevity-linked securities are investigated. As the longevity market is an immature market, based on a non-financial risk, the classical methodology of risk-neutral pricing cannot be used carelessly. Indeed, the lack of liquidity in the market induces incompleteness, as it is typically the case when non-hedgeable and non-tradable claims exist. Thus, to price some financial contracts on longevity-risk, a classical arbitrage-free pricing methodology is inapplicable as it relies upon the idea of risk replication. The replication technique is only possible for markets with high liquidity and for deeply traded assets. It induces a unique price for the contingent claim, which is the cost of the replicating portfolio hedging away the market risk. Hence, in a complete market, the price of the contingent claim is the expected future discounted cash-flows, calculated by the unique risk-neutral probability measure. In contrast, in an incomplete market, such as a longevity-linked securities market, there will be no universal pricing probability measure, making the choice of pricing probability measure crucial.

What will be a good pricing measure for longevity ? It is expected that the historical probability measure will play a key role, due to the reliable data associated with it. Therefore, it

8. This bond project with a maturity of 25 years included coupons based on a survivor index

seems natural to look for a pricing probability measure equivalent to the historical probability measure. Important factors to consider, are that a relevant pricing measure must be : robust with respect to the statistical data, but also compatible with the prices of the liquid assets quoted in the market. Therefore, a relevant probability measure should make the link between the historical vision and the market vision. Once the subsets of all such probability measures that capture the desired information is specified, a search can commence for the optimal example by maximising the likelihood or the entropic criterion.

Characterisation of a pricing rule

Questions concerning the pricing rule are also essential when addressing financial transactions in an incomplete market. If a risky cash flow F with a long-term maturity is considered ; for example, for the EIB-BNP Paribas longevity bond mentioned in Section 4.4.5, the cash flow F corresponds to the sum of the discounted coupons.

Classical static pricing methods The change from the historical measure to the pricing measure introduces a longevity risk premium. This method is similar to those based upon actuarial arguments, where the price of a particular risky cash-flow F can be obtained as

$$\pi(F) = \mathbb{E}_{\mathbb{P}}(F) + \lambda \sigma_{\mathbb{P}}(F).$$

The risk premium λ is a measure of the Sharpe ratio of the risky cash-flow F . Different authors have studied the impacts of different choices of probability measures on the pricing ([34]) for the other types of financial contracts, such as weather derivatives.

As recalled before, in a highly liquid and complete market where risky derivatives can be replicated by a self-financing portfolio, the risk-neutral (universal) pricing rule is used :

$$\pi(F) = \mathbb{E}_{\mathbb{P}^*}(F) = \mathbb{E}_{\mathbb{P}}(F) + \text{cov}\left(F, \frac{d\mathbb{P}^*}{d\mathbb{P}}\right).$$

This pricing rule is linear, as the actuarial rule does not take into account the risk induced by large transactions. However, in those cases where hedging strategies cannot be constructed, the nominal amount of the transactions becomes an important risk factor. In such cases, this methodology is no longer accurate, especially when the market is highly illiquid. A seemingly more appropriate methodology to address this problem is the utility based indifference pricing methodology presented below.

Indifference pricing In an incomplete market framework, where perfect replication is no longer possible, a more appropriate strategy involves utility maximisation. Following [33], the maximum price that an agent is ready to pay is relative to their indifference towards the transaction and according to their individual preference. More precisely, given

a utility function u^b and an initial wealth of W_0^b , the indifference buyer price of F is $\pi^b(F)$ determined by the non-linear relationship :

$$\mathbb{E}_{\mathbb{P}}(u^b(W_0^b + F - \pi^b(F))) = \mathbb{E}_{\mathbb{P}}(u^b(W_0^b)).$$

This price, which theoretically depends upon initial wealth and the utility function, is not necessarily the price at which the transaction will take place. This specifies an upper bound to the price the agent is ready to pay. Similarly, the indifference seller price is determined by the preference of the seller, and characterised by

$$\mathbb{E}_{\mathbb{P}}(u^s(W_0^s - F + \pi^s(F))) = \mathbb{E}_{\mathbb{P}}(u^s(W_0^s)). \quad (4.2)$$

It should be noted that this pricing rule is non-linear and difficult to compute, and provides a price-range rather than a single price.

Fair price for small transactions Agents who are aware of their sensitivity to unhedgeable risk can attempt to transact at a low price when forming a risky contract. In this case, the buyer wants to transact at the buyer's "fair price" : which corresponds to the zero marginal rate of substitution p^b .

$$\begin{aligned} \partial_{\theta} \mathbb{E}_{\mathbb{P}}(u^b(W_0^b + \theta F))|_{\theta=0} &= \partial_{\theta} \mathbb{E}_{\mathbb{P}}[u^b(W_0^b - \theta p(F))]|_{\theta=0} \\ \text{that implies } p^b(F) &= \mathbb{E}_{\mathbb{P}}(u_x^b(W_0^b)F) / \mathbb{E}_{\mathbb{P}}(u_x^b(W_0^b)). \end{aligned} \quad (4.3)$$

The same formula holds when the random initial wealth W_0^b is in fact the value of the optimal portfolio found in the classical optimisation problem when considering the scenario of an incomplete market utility function u^b . In this case, the normalised random variable $u_x^b(W_0^b)$ may be viewed as the optimal martingale measure.

When both agents have the same utility function, they can transact at this fair price. Therefore, with an exponential utility function, the fair price for a small transaction is equivalent to the one given by the expectation under an equivalent probability measure. This methodology can be compared to the Wang transform, which is a distortion of the historical probability (see the approach of [17]). If we linearise the Wang transform an equivalent change in probability measures can be produced. The potential shortcomings of this approach are discussed in [4].

Economic point of view When approaching the same subject from an economic point of view, the transaction price can be said to form an equilibrium price. This occurs between either the seller and buyer, or different players in the market. It can be described as a transaction where the agents simultaneously maximise their expected utility (Pareto-optimality). Obviously a transaction only takes place where there are two agents for whom $\pi^b(F) \geq \pi^s(F)$.

A dynamic point of view

The economic pricing methodologies described in the above subsection are static and correspond in this respect mainly to an insurance or accounting point of view. However, the standard financial approach to pricing is different, as it relies on the so-called, risk-neutral methodology. The main underlying assumption of this approach is that it is possible to replicate cash flows of a given transaction dynamically using basic traded securities in a highly liquid market. Using a non-arbitrage argument, the price of the contract is uniquely defined by the cost of the replicating strategy. Using a risk-neutral probability measure as a reference, it can also be proven that this cost is in fact the expected value of the discounted future cash flows. This approach is clearly dynamic, since the replicating strategy is dynamically constructed. Note : the replicating portfolio is not only a tool to find the price of the contract, but can also be used to dynamically hedge the risks associated with the transaction. Since the cash flows are discounted and longevity-linked securities have a long maturity, this approach raises some specific issues relating to dynamic long-term interest-rates (see subsection 4.5.2). In any case, adopting such an approach for the pricing of financial contracts based upon mortality- or longevity-risks requires the underlying risks to be dynamically modelled (see subsection 4.5.2). It is important to emphasise that a highly liquid underlying market is essential for the construction of such a replicating strategy. Nevertheless, as has been mentioned, the present state of the longevity market is far from liquid, and the applicability of risk neutral methodology has been questioned in many research papers ; including [4]. It could be argued that to extend the fair price approach to an illiquid and dynamic setting is an appropriate action. Since a perfect hedge does not exist, we can extend (4.3) as

$$\sup \left(\mathbb{E}_{\mathbb{P}} \left(u^s(W_0^s - F + \pi^s(F)) \right) \right) = \mathbb{E}_{\mathbb{P}} \left(u^s(W_0^s) \right), \quad (4.4)$$

where $\pi^s(F)$ is no longer a static price, but a dynamic price strategy associated with a hedge, and the supremum in (4.4) is taken over the strategies. Thus an optimal hedge, given criteria (4.4) is derived (see [3]).

Design issues

Due to the absence of liquidity and maturity in the longevity market, a dynamic replicating strategy cannot be constructed ; and therefore, the various risks embedded in a longevity transaction will be difficult to dynamically hedge . As a consequence, investors will consider the pricing and design of a transaction in order to select the one, which from their point of view seems the least risky. Therefore, the design of new securities is clearly an extremely important feature of the transaction ; in fact, it may mean the difference between success and failure.

The problem that will be considered next, is the design of a contract F and a price $\pi(F)$ that are desirable for both the buyer and seller to form a transaction. Such a problem

is studied in detail by [3]. An interesting feature of the risk transfer emerges when both agents have the same utility function but different risk tolerances : $u^b(x) = \gamma^b u(x/\gamma^b)$ and $u^s(x) = \gamma^s u(x/\gamma^s)$. Given an initial risk that the seller wants to partially transfer to the buyer, the best choice is to transfer $F^* = \frac{\gamma^b}{\gamma^s + \gamma^b} X$. Thanks to the optimality of F^* , it can be shown that a pricing probability measure common to both agents still exists, but that it depends upon the pay-out of the derivatives. Nevertheless, this feature is in some ways simplistic, as investors do not usually share the same utility function (because, for example, they do not have access at the same market). Thus, such issues are not reduced in the transfer mentioned above ; but, in fact, become more crucial and complex.

4.5.2 Dynamic modelling of the underlying risk

Dynamic pricing approaches raise a number of issues related to the modelling of the underlying risk. In addition to longevity-risk, financial contracts on longevity are also exposed to long-term interest rate risk. Due to the lack of liquidity on the long-term horizon, the interest rate risk should be handled differently from the rate explored in typical financial modelling. In addition, as introduced in Section 4.2, the underlying longevity-risk should be undertaken through a continuous-time rather than a discrete-time framework.

Long-term interest rates

Financial contracts on mortality-related risks, typically have a maturity up to 20 years. Whilst in contrast, longevity-linked securities are typically characterised by a much longer maturity (40 years and beyond). In the case of the vast majority of such contracts, there is an embedded interest rate risk. For the shorter-term time-horizon, the standard financial models perform appropriately to model the underlying risk. However, this is no longer the case for longer maturities. In such scenarios, the interest rate market becomes highly illiquid and the typical financial attitude is not easily extendable. Nevertheless, an abundant literature on the economic aspects of long-term policy-making has been developed (see for example [25, 26], [32] or [9]). Much of this discourse focuses on the aggregate behaviour of all agents, and the representation of the economy is limited to the strategy of the representative agent. The derivation of the yield-curve for far-distant maturities is induced from the maximisation of the representative agent's inter-temporal utility function on the aggregate consumption. This leads to the so-called "Ramsey rule". Nevertheless, in the economic literature on long-term policy-making, dynamic modelling processes are not common, with the majority of space dedicated to deterministic methodology with static random perturbations. Furthermore, the global approach with a representative agent never takes into account the existence of a financial market. To price and study longevity-linked securities involving clearly long-term interest rates, the financial markets need to be put at the heart of methodological preoccupations (see a forthcoming paper of [14]).

Stochastic mortality modelling for financial transactions

For the effective pricing longevity linked securities, future mortality forecasting models need to be developed that are more accurate than those currently used by insurance companies. Thus, the models introduced in Section 4.2.1 can be enriched by factoring in more uncertainty into their parameters. As previously discussed, in a deterministic approach for the hazard rate $\mu(x, t)$, the survival probability of an individual aged x at time t is defined as :

$$S_t(x, T) = \mathbb{P}_t(\tau > x + (T - t) \mid \tau > x) = \exp\left(-\int_t^T \mu(x + s - t, s) ds\right). \quad (4.5)$$

This can also be extended to a stochastic hazard rate $\mu(x, t)$ using two different approaches : Firstly, by analogy with periodic life tables, the period approach models the process $s \rightarrow \mu(x, s)$ for each age x ; in this case, $\mu(x, s)$ is referred to as the spot mortality rate (see [40]). Secondly, by analogy with life tables, the approach is cohort-based, and instead, models the process $s \rightarrow \mu(x + s, s)$, taking into account the evolutionary future of mortality.

More precisely, let $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a probability space representing the uncertainty under the historical probability measure \mathbb{P} . The intensity process $\mu(x, t)$ is adapted to the information \mathcal{F}_t available to the agents at any time t . The survival probability is now defined as :

$$S_t(x, T) = \mathbb{P}_t(\tau > x + (T - t) \mid \tau > x) = \mathbb{E}_t\left(\exp\left(-\int_t^T \mu(x + s - t, s) ds\right)\right). \quad (4.6)$$

Here $\mathbb{P}_t(\cdot)$ stands for the conditional probability of \mathbb{P} with respect to \mathcal{F}_t , and x for the age of the cohort at time t .

Using Equation (4.6), the hazard rate can be defined as :

$$\begin{aligned} \mu(x, T) &= -\lim_{t \uparrow T} \partial_T \ln S_t(x, T) \\ &= \lim_{t \uparrow T} \left[-\frac{\partial}{\partial T} \ln \mathbb{E}_t\left(\exp\left(-\int_t^T \mu(x + s - t, s) ds\right)\right) \right]. \end{aligned} \quad (4.7)$$

Note that $\mu(x, T)$ is indeed $\mu(x_T, T)$, where x_T is the age of the individual at time T .

Alternatively, if a cohort based mortality model is considered, the quantity can be defined as $\mu_t(x, T)$, where x is the age at time t by :

$$\mu_t(x, T) = -\frac{\partial}{\partial T} \ln \mathbb{E}_t\left(\exp\left(-\int_t^T \mu(x + s - t, s) ds\right)\right). \quad (4.8)$$

By analogy with interest rate frameworks or credit risk modelling, $\mu_t(x, T)$ is termed as the instantaneous forward mortality rate. Note that $S_t(x, T)$ in this case is equivalent to the price at time t of a defaultable zero-coupon bond maturing at T . This definition is

extremely useful when valuing insurance contracts, as it takes into account the whole age-term structure of mortality, which, is consistent with insurer practices.

Yet, in the periodic approach of mortality modelling (*spot* mortality rates), some authors ([42] and [20] or [8]) introduce the so-called intensity framework to stochastic mortality modelling, which is very popular both in interest rate and credit risk modelling. By doing so, they draw an analogy between death-time and the default-time. Moreover, this methodology underpins the derivation of closed-form formula for survival probabilities, especially when assuming an affine diffusion for the mortality hazard rate process $\mu(x, t)$. However, this class of processes shows restrictive behaviours, and does not ensure positive mortality intensities. In fact, [29] have outlined these drawbacks, introducing a new class of processes : the so-called quadratic class. Following this approach, [28] use a quadratic stochastic intensity model with a Gaussian autoregressive factor to model data on spot mortality rates; a method proving to be more suitable and consistent with historical data. They also emphasise the tractability of such processes when used to evaluate insurance contracts that apply a risk-neutral approach. Nevertheless, the intensity approach for mortality is more complex than in credit-risk, because the hazard rate depends on the age x , which is a difficult dependency to model and capture in such a framework.

Consequently, it is essential from this point on to understand precisely what data the various parameters of the models have to estimate statistically. The most accurate and reliable data available is the national data of mortality. However, as we have mentioned in Section 2.4, these mortality tables are heterogeneous and present significant differences from one country to another.

The statistical data will be used to construct pricing methodologies for longevity-linked securities. For derivatives written on non-financial risk in an immature market, there would be no financial benchmark to estimate the market risk premium.

4.6 Conclusion

This paper has discussed the most up-to-date knowledge and modelling on longevity risk. The specificities of longevity risk compared with other insurance risks - concerning in particular, trend sensitivity, geographical variability and the associated long-term maturities - means longevity risk is highly intricate and all the more so considering the potential financial and non-financial correlations with other risk sources.

Recently, the life insurance sector and insurance industry in general have been set to adopt new regulations intended to ensure increased accuracy in risk-assessment, and tighter rules governing the management of solvency risk. These regulations, combined with the increasing convergence between the insurance and capital markets, has opened the way for alternative risk management solutions and innovative risk transfers. To support the growth of this emerging market, a number of actions must be taken. Firstly, asymmetry of information

between the various agents potentially involved in transactions needs to be reduced. Also, it is necessary to develop specific pricing methods and partial hedging methodologies that are better-suited to both the particular features of longevity-risk and to the immature and illiquid state of the current market. However, as a result of the long maturities of the potential transactions, additional challenges are also expected for the modelling of long term interest rates.

The challenges of longevity-risk are not only a concern for the insurance industry ; they are at the very core of the political, economic and strategic discourses that will guide the the design of the "effective" retirement age policies and pension schemes that will meet the future challenges of our society. The potential impacts of longevity risk are highly significant at various levels of society and the global economy, which will make better management of these risks one of the key challenges of the coming decade. Interactions with other sources of risk, such as dependence, economy and ecology, have not been investigated in-depth in this paper : however, it would be of great interest to see future research that studies the impacts of longevity on the larger economy and the environment at the macro-level.

Bibliographie

- [1] Antolin P. and H. Blommestein (2007), “Governments and the market for longevity-indexed bonds”, *OECD Journal : Financial Market Trends*, 4, 92.
- [2] Barrieu P. and L. Albertini (2009), “The handbook of insurance-linked securities”, Wiley Finance.
- [3] Barrieu P. and N. El Karoui (2009), “Pricing, hedging and optimally designing derivatives via minimization of risk measures”, *Volume on Indifference Pricing, Princeton University Press*, 77-146.
- [4] Bauer D., M. Boerger and J. Russ (2010), “On the Pricing of Longevity-Linked Securities”, *Insurance, Mathematics and Economics*, 46, 1, 139-149.
- [5] Bensusan H. and N. El Karoui (2009), “Microscopic Models for Populations Trait and Age Structures and Applications for Longevity problem”, *Working Paper, available at [http :www.cmap.polytechnique.fr/bensusan](http://www.cmap.polytechnique.fr/bensusan)*.
- [6] Barrieu P., H. Bensusan, N. El Karoui, S. Loisel, C. Ravanelli, Y. Salhi (2010), “Understanding, Modelling and Managing Longevity Risk : Key Issues and Main Challenges”, *Working Paper, available at [http :www.cmap.polytechnique.fr/bensusan](http://www.cmap.polytechnique.fr/bensusan)*.
- [7] Bienvenue A., A. Illig, S. Loisel and D. Serant (2010), “On inter-age correlations in stochastic mortality models”, *Tech. rep.*
- [8] Biffis E. (2005), “Affine processes for dynamic mortality and actuarial valuations”, *Insurance : Mathematics and Economics*, 37, 3, 443-468.
- [9] Breeden D.T. (1979), “An intertemporal asset pricing model with stochastic consumption and investment opportunities”, *Journal of Financial Economics*, 7, 3, 265-296.
- [10] Brouhns N. and M. Denuit (2002), “Risque de longévité et rentes viagères. Tables de mortalité prospectives pour la population belge”, *Belgian Actuarial Bulletin*, 2, 1.
- [11] Blake D., A. Cairns and K. Dowd (2006), “A two-factor model for stochastic mortality with parameter uncertainty : Theory and calibration”, *Journal of Risk and Insurance*, 73, 687-718
- [12] Blake D., A. Cairns and K. Dowd (2008), “Modelling and management of mortality risk : a review”, *Scandinavian Actuarial Journal*, 2, 79-113.

- [13] Blake D., A. Cairns, K. Dowd, G.D. Coughlan, D. Epstein, A. Ong and I. Balevich (2007), “A quantitative comparison of stochastic mortality models using data from England & Wales and the United States”, Working paper, Heriot-Watt University, and Pensions Institute Discussion Paper PI-0701.
- [14] Camilier I., N. El Karoui and C. Hillairet (2010), “Long term interest rate and Growth optimal portfolio”, *Working Paper*
- [15] Ceiops (2008), “Quantitative impact study.”, *Tech. rep., European Commission, Available from <http://www.ceiops.eu/content/view/118/124/>*
- [16] Cox S.H and Y. Lin (2008), “Securitization of catastrophe mortality risks”, *Insurance : Mathematics and Economics*, 42, 2, 628-637.
- [17] Cox S.H. and Y. Lin (2007), “Natural hedging of life and annuity mortality risks”, *North American Actuarial Journal*, 11, 3, 1-15.
- [18] Currie I.D., M. Durban and P.H.C Eilers (2004), “Smoothing and forecasting mortality rates”, *Statistical Modelling*, 4, 4, 279-298.
- [19] Currie I.D., J. Kirkby (2010), “Smooth models of mortality with period shocks”, *Statistical Modelling*, 10, 2, 177-196.
- [20] Dahl M. (2004), “Stochastic mortality in life insurance : market reserves and mortality-linked insurance contracts”, *Insurance : Mathematics and Economics*, 35, 1, 113-136.
- [21] Delwarde A., D. Kachakhidze, L. Olié and M. Denuit (2004), “Modèles linéaires et additifs généralistes, maximum de vraisemblance local et méthodes relationnelles en assurance sur la vie”, *Bulletin Français d’Actuariat*, 6, 77-102.
- [22] Filipovic D. (2009), “Multi-Level Risk Aggregation”, *ASTIN Bulletin*, 39, 2, 565-575.
- [23] FAP (2007), “Personnes âgées et médicaments”, *Working Paper, Available from <http://www.acadpharm.org/medias/direct/Agees.pdf>*
- [24] Fournier N. and S. Méléard (2004), “A microscopic probabilistic description of a locally regulated population and macroscopic approximations”, *Annals of Applied Probability*, 14, 4, 1880-1919.
- [25] Gollier C. (2007), “The consumption-based determinants of the term structure of discount rates”, *Mathematics and Financial Economics*, 1, 2, 81-101.
- [26] Gollier C. (2008), “Discounting with fat-tailed economic growth”, *Journal of Risk and Uncertainty*, 37, 2, 1-16.
- [27] Goshay R.C. and R. Sandor (1973), “An inquiry into the feasibility of a reinsurance futures market”, *Journal of Business Finance*, 5, 2, 56-66.
- [28] Gourieroux C. and A. Monfort (2008), “Quadratic stochastic intensity and prospective mortality tables”, *Insurance : Mathematics and Economics*, 43, 1, 174-184.
- [29] Gourieroux C. and R. Sufana (2003), “Wishart quadratic term structure models”, *Tech. rep., Les Cahiers du CREF of HEC Montreal Working Paper*, 3, 10.

-
- [30] Haberman S. and A.E. Renshaw (2003), "On the forecasting of mortality reduction factors", *Insurance : Mathematics and Economics*, 32, 3, 379-401.
- [31] Haberman S. and A.E. Renshaw (2006), "A cohort-based extension to the Lee-Carter model for mortality reduction factors", *Insurance : Mathematics and Economics*, 38, 3, 556-570.
- [32] Hansen L.P. and J.A. Scheinkman (2009), "Long-term risk : An operator approach", *Econometrica*, 77, 1, 177-234.
- [33] Hodges S.D. and A. Nauberger (1989), "Optimal replication of contingent claims under transaction costs", *Review of Futures Markets*, 8, 2, 222-239.
- [34] Jewson S. and A. Brix (2005), "Weather derivative valuation : the meteorological, statistical, financial and mathematical foundations", *Cambridge University Press*.
- [35] Jusot F. (2004), "Mortalité, revenu et inégalités de revenu en France : une analyse multi-niveaux", IRDES, Reference BDSP : 309602.
- [36] Klein J.P. and M.L. Moeschberger (2003), "Survival analysis : techniques for censored and truncated data", *Springer, Verlag*.
- [37] Lee R.D. and L.R. Carter (1992), "Modeling and forecasting U.S. mortality", *Journal of the American Statistical Association*, 87, 659-675.
- [38] Loisel S. and D. Serant (2007), "In the core of longevity risk : hidden dependence in stochastic mortality models and cut-offs in prices of longevity swaps", *Insurance : Mathematics and Economics*.
- [39] Loisel S. and Y. Salhi (2010), "Joint modeling of portfolio experienced and national mortality : A co-integration based approach", *Working Paper, Available from <http://isfa.univ-lyon1.fr/stephane.loisel/>*.
- [40] Milevsky M.A. and S.D. Promislow (2001), "Mortality derivatives and the option to annuities", *Insurance : Mathematics and Economics*, 29, 3, 299-318.
- [41] Pitacco E. (2004), "Survival models in a dynamic context : a survey", *Insurance : Mathematics and Economics*, 35, 2, 279-298.
- [42] Schrager D.F. (2006), "Affine Stochastic mortality", *Insurance : Mathematics and Economics*, 38, 1, 81-97.
- [43] Tran V.C. (2006), "Modèles particuliers stochastiques pour des problèmes d'évolution adaptative et pour l'approximation de solutions statistiques", Dissertation.
- [44] Tran V.C. (2008), "Large population limit and time behaviour of a stochastic particle model describing an age-structured population", *ESAIM : P&S*, Vol 12, 345-386.

BIBLIOGRAPHIE

Chapitre 5

Modélisation de la mortalité individuelle

Introduction

Ce chapitre concerne l'étude d'un modèle de mortalité individuel par âge et par traits. Nous caractérisons l'influence de certaines caractéristiques sur la mortalité et nous étudions la calibration du modèle.

Motivations

L'évolution de la mortalité dans le temps est assez complexe et de nombreux chercheurs essaient de la modéliser. Historiquement, le logarithme du taux de mortalité des individus évolue dans le temps avec une tendance et des variations autour de cette tendance. Les modèles mathématiques standards de mortalité essaient de capter la tendance et la volatilité de cette évolution. Les modèles de type Cairns Dowd and Black (CDB) [3, 4] ou les modèles de type Lee Carter [19, 20, 27] considèrent que le taux de mortalité, pour un sexe donné, est une fonction de l'âge et du temps uniquement. Ces modèles peuvent également capter un effet cohorte ou un effet générationnel [37, 29, 28] en rajoutant une composante de mortalité dépendant de la date de naissance de l'individu.

Il existe d'autres type de modèles de mortalité stochastique dont la plupart sont inspirés de modèles utilisés dans le monde de la finance. Les modèles de mortalité avec une diffusion de type affine [33, 15] s'inspirent du monde des taux d'intérêt et permettent de décrire au mieux ces effets de tendance et de volatilité. De plus, des ruptures de tendance sont observées dans l'évolution de la mortalité et apparaissent espacées dans le temps à des intervalles de temps de quelques dizaines d'années. Des travaux récents cherchent à capter ces ruptures de tendance de mortalité dans le temps afin d'en tenir compte dans les projections futures

à des horizons lointains [14].

Afin d’avoir une meilleure description de la mortalité, il serait intéressant de déterminer un niveau mortalité spécifique à chaque individu. Actuellement, les informations utilisées sont l’âge et le sexe mais il existe d’autres variables (individuelles et environnementales) qui ont un impact significatif. Les tables de mortalité nationale correspondent à un niveau moyen de mortalité et nous cherchons à avoir un taux de mortalité individuel qui tient compte des caractéristiques des individus.

Une application importante de ce type de modèle est le domaine de l’assurance vie. De nouvelles tables de mortalité par âge et par traits pourraient être très utiles dans le monde de l’assurance. En effet, une grande problématique des assureurs est que la mortalité des assurés est différente de la mortalité moyenne nationale. Les assurés vivent intrinsèquement plus longtemps puisqu’ils ont a priori un statut socioéconomique supérieur à la moyenne. Les assureurs observent cet écart significatif de mortalité mais ont du mal à le quantifier et à en tenir compte : ce phénomène est appelé risque de base ou risque spécifique. Les modèles standards de mortalité par âge n’aborde uniquement que le risque de longévité systémique (risque à l’échelle nationale). L’extension à ce type de modèle permettrait de prendre en compte à la fois le risque systémique et le risque spécifique.

Présentation du travail

Ce chapitre considère un modèle de mortalité qui s’inspire des modèles de type Cairns tenant compte non seulement de l’âge et du temps mais aussi d’un ensemble de caractéristiques susceptibles de modifier le taux de mortalité d’un individu. L’idée de ce modèle est de tenir compte de toute l’information disponible sur les individus pour évaluer précisément leur taux de mortalité.

Selon certains démographes, la réduction de la mortalité en Europe résulterait pour 20% de l’évolution du niveau socioéconomique des individus, pour 30% de l’évolution de l’éducation et des comportements et pour 50% des nouvelles connaissances médicales et technologiques [26]. Nous pouvons d’ailleurs citer des études économiques effectuées par l’IRDES [17, 8] qui décrivent l’existence d’une relation entre le niveau socioéconomique et la mortalité : par exemple, en ce qui concerne l’espérance de vie des individus de 35 ans, il y a actuellement un écart de plus de 6 ans entre les cadres et les ouvriers. De plus, les conditions de vie liées à l’environnement dans lequel un individu évolue ont elles aussi un impact significatif sur son niveau de mortalité. L’analyse de ces études permet de sélectionner les traits qui vont impacter significativement les taux de mortalité et d’avoir une intuition sur l’influence de chaque caractéristique. Le modèle de mortalité par âge et par traits prendra en compte ces caractéristiques et nous vérifierons que ce modèle confirme la plupart de ces intuitions.

Enfin, avec l'émergence de nouvelles pandémies de grippe dans le monde, il nous est apparu important de tenir compte de la surmortalité due aux catastrophes naturelles ou humaines. Au cours du vingtième siècle, l'Europe a connu 3 grandes catastrophes comme la grippe espagnole qui a fait des millions de morts ainsi que les deux guerres mondiales. La prise en compte de ce type d'événements dans la modélisation a un impact important sur l'étude de la mortalité à long terme.

Tout d'abord, nous rappelons le modèle de Cairns classique par âge dont nous détaillons les propriétés. Nous décrivons la méthode de calibration de ce modèle et nous proposons une étude empirique de "backtesting" pour sélectionner la plage de données à utiliser pour la calibration. Nous proposons une méthode de la prise en compte des pandémies et catastrophes naturelles dans la modélisation du taux de mortalité.

Ensuite, nous analysons différentes études économiques et démographiques décrivant une étude statistique de la mortalité. Ces études permettent d'identifier quelles sont les caractéristiques, individuelles (autres que l'âge) et globales (liées au pays ou à la région) ayant une influence sur le niveau de mortalité. Nous faisons référence à de nombreuses études économiques et démographiques afin de déterminer les facteurs explicatifs de mortalité.

Dans une troisième partie, nous introduisons un modèle dans lequel le taux de mortalité dépendra de certaines caractéristiques individuelles et de certaines caractéristiques globales de l'environnement dans lequel la population évolue. L'idée est que le modèle doit rester cohérent avec les modèles classiques de mortalité par âge calibrés sur les tables de référence qui décrivent un taux de mortalité national en fonction de l'âge et du sexe (Human Mortality Database (HMD), INSEE). Ainsi, lorsqu'une information sur l'individu est disponible comme par exemple le statut matrimonial, son taux de mortalité est modifié. Cependant, si cette information n'est pas disponible, il faut que le taux de mortalité soit égal au taux de mortalité national par âge et par sexe.

Afin de calibrer ce modèle, nous utilisons des données issues de l'échantillon démographique permanent¹ (EDP) de l'INSEE [16] qui fournissent un taux de mortalité par âge et par traits. L'EDP est un échantillon de la population française dont l'effectif correspond à environ 1% de la population totale et pour lequel chaque individu est recensé avec un ensemble de caractéristiques (catégories socioprofessionnelles, statut matrimonial, niveau de diplômes,...). Nous proposons une méthode de calibration du modèle de mortalité par âge et par traits sur ces données individuelles.

1. Nous remercions l'INSEE et en particulier Antoine Frachot et Xavier Niel pour nous avoir autorisé à utiliser ces données précieuses

Finalement, nous avons étudié l'impact de la catégorie socioprofessionnelle, du niveau d'éducation et du statut matrimonial pour lequel nous avons pu nous procurer des données. Ce modèle prend aussi en compte l'évolution de l'influence de ces facteurs dans le temps. En effet, l'évolution des mœurs et le changement des comportements modifient l'influence de certains traits au cours du temps. Par exemple, la notion de mariage a fortement évolué ces dernières années avec un nombre important de couples vivant ensemble sans être mariés. L'extension de la notion classique de taux de mortalité par âge à un taux de mortalité individuel qui dépend de l'âge et de traits caractéristiques des individus et de l'environnement du pays permettrait de créer de nouvelles tables prospectives de mortalité.

5.1 Étude du modèle de Cairns décrivant la mortalité par âge

Il existe de nombreux modèles de projections de mortalité comme le modèle de Lee Carter [19, 20, 27], le modèle de Cairns [3, 4, 5] et les modèles affines [33, 15] pour n'en citer que quelques uns. Il aurait été possible de mener l'étude qui suit avec un autre modèle de mortalité tel que le modèle de Lee Carter. Nous nous intéressons plus précisément au modèle classique de mortalité à deux facteurs introduit par Cairns (M5). Ce modèle consiste en une régression logistique du taux de mortalité par rapport à l'âge de l'individu (avec distinction éventuelle du sexe) en considérant que les coefficients de la régression sur l'âge sont stochastiques. Le modèle de Cairns va présenter de nombreux avantages pour l'étude menée dans la suite dans la thèse (Chapitre 6, Chapitre 7, Chapitre 8). En effet, le taux de mortalité reste rigoureusement borné entre 0 et 1 grâce à l'utilisation de la fonction "logit" et les variables catégorielles ont une interprétation intéressante en terme de régression logistique.

5.1.1 Description et justification du modèle de mortalité

Le taux de mortalité pour l'âge a à la date t est noté $q(a, t)$ correspond à la probabilité d'un individu d'âge a de décéder au cours de l'année calendaire t . Le modèle de Cairns classique modélisant la fonction $\text{logit}(q(a, t)) = \ln\left(\frac{q(a, t)}{1-q(a, t)}\right)$ à l'aide d'un facteur $A(t) = (A_1(t), A_2(t))^T$ s'écrit de la manière suivante :

$$\begin{aligned} \text{logit}[q(a, t)] &= A_1(t) + aA_2(t) \\ A(t+1) &= A(t) + \mu + CZ(t+1), \end{aligned} \tag{5.1}$$

où $\mu = (\mu_1, \mu_2)^T$ est le drift représenté par un vecteur constant de taille 2×1 et C est une matrice 2×2 triangulaire inférieure de Cholesky reflétant les volatilités et la corrélation des facteurs A_1 et A_2 . Les composantes du vecteurs gaussien $Z(t)$ sont des variables gaussiennes

indépendantes centrées réduites. L'estimation des paramètres μ et C se fait par régression linéaire sur l'historique des courbes de mortalité par âge. En effet, à chaque date t , on considère la courbe des logit de mortalité par âge $\text{logit}(q(t, a))$ et on calcule les coefficients de la régression linéaire $A_1(t)$ et $A_2(t)$. Ce procédé fournit un échantillon de coefficients dont les différentiels sont supposés gaussiens de moyenne μ et de matrice de variance-covariance $\Sigma = C^T C$. Ce modèle inclut une composante d'amélioration de la longévité $A_1(t)$ indépendante de l'âge et une deuxième composante $A_2(t)$ qui mitige les effets d'amélioration aux âges élevés. Il existe éventuellement une corrélation entre ces deux facteurs au travers de la matrice C , mais cette corrélation n'aura pas une influence majeure étant donné son ordre de grandeur. La dynamique des facteurs $A_1(t)$ et $A_2(t)$ dépend de paramètres constants et décrit la mortalité à l'échelle nationale.

Dans le modèle classique de Cairns, le "logit" de mortalité est modélisé comme une fonction affine de l'âge :

$$\text{logit}(q(a, t)) = \text{logit}(q(a, t-1)) + \mu(a) + \sigma(a)Z_t,$$

où Z_t est une loi normale centrée réduite tirée à la date t et où $\sigma(a)^2 = (C_{11} + aC_{21})^2 + a^2C_{22}$.

En analysant l'historique des logit de courbes de mortalité, on observe un caractère linéaire en temps avec une tendance qui dépend de l'âge ainsi qu'un comportement variationnel autour de cette tendance. En dehors des deux grandes guerres, l'effet tendance en France est justifié et permet d'expliquer pour une grande part l'évolution de la mortalité pour la tranche d'âge 30/90 ans. On représente les courbes de mortalité des hommes (il en aurait été de même pour les femmes/ pour certains âges entre 1900 et 2007 et on observe effectivement une évolution linéaire dans le temps avec des variations autour d'une tendance. On note aussi la présence de ruptures de pentes au cours du siècle dernier (voir Figure 5.1 à Figure 5.4)

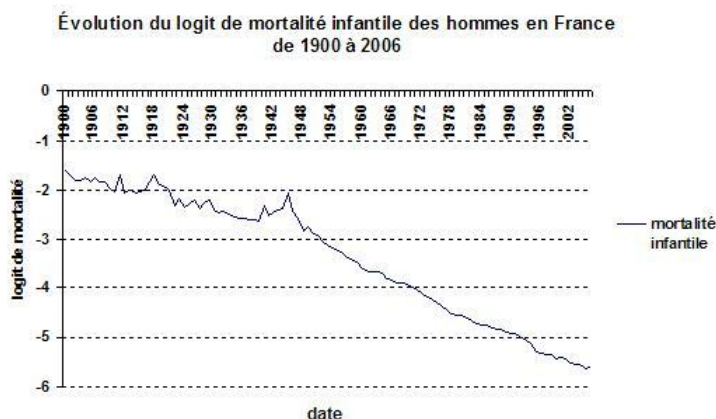


FIGURE 5.1 – Évolution du logit de la mortalité infantile des hommes en France

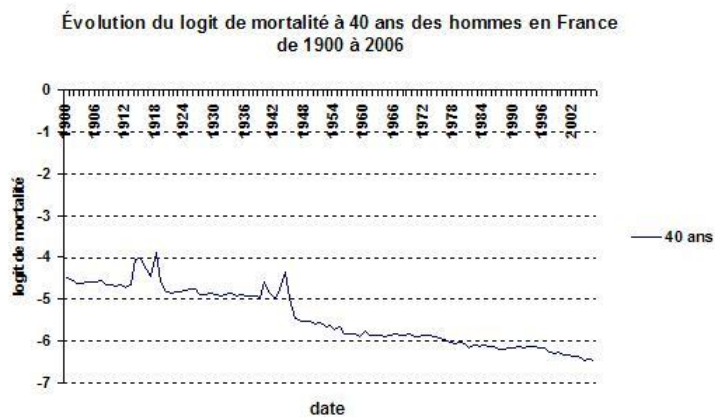


FIGURE 5.2 – Évolution du logit de la mortalité des hommes de 40 ans en France

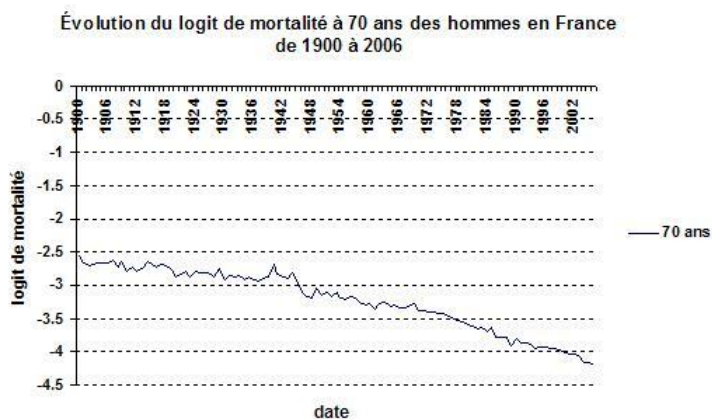


FIGURE 5.3 – Évolution du logit de la mortalité des hommes de 70 ans en France

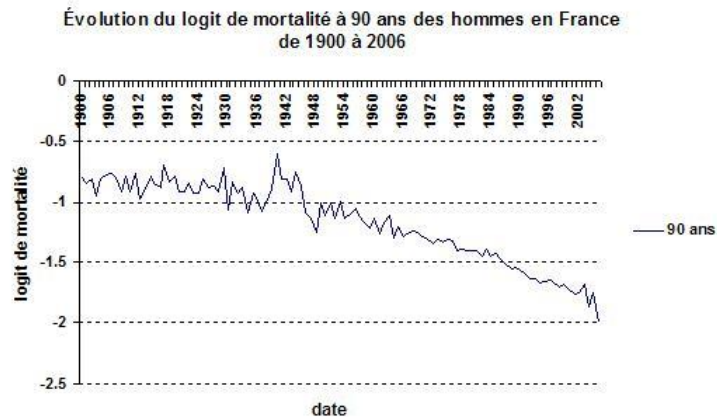


FIGURE 5.4 – Évolution du logit de la mortalité des hommes de 90 ans en France

De plus, le modèle consiste à considérer un drift et une volatilité linéaire en l'âge et ceci se justifie encore largement en regardant l'historique des courbes de mortalité en fonction de l'âge à certaines périodes. En effet, à partir de l'âge de 35 ans, la linéarité en âge est presque parfaite et on peut éventuellement capter la légère non linéarité en effectuant une régression quadratique (voir Figure 5.5 à Figure 5.7). À partir de 1990, la non-linéarité en âge s'intensifie et il faudra penser à améliorer le modèle pour les âges faibles dans le but d'une modélisation de la démographie.

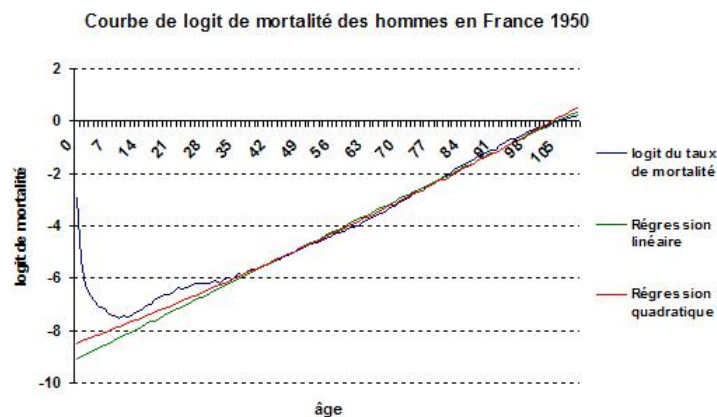


FIGURE 5.5 – Évolution du logit de la mortalité France en fonction de l'âge en 1950

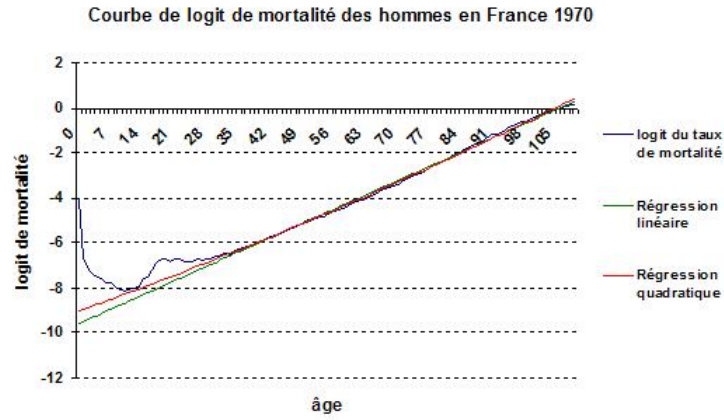


FIGURE 5.6 – Évolution du logit de la mortalité France en fonction de l'âge en 1970

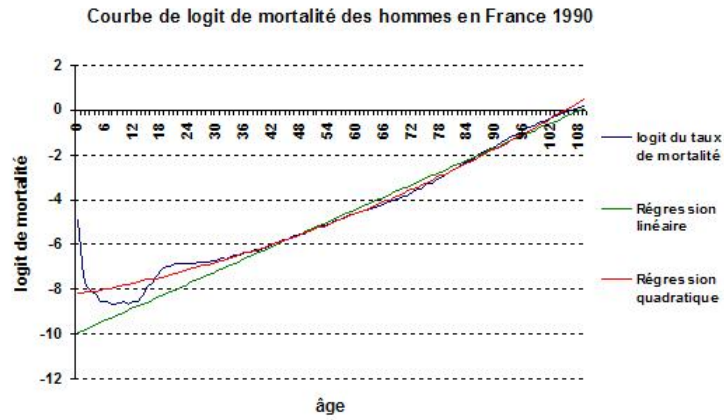


FIGURE 5.7 – Évolution du logit de la mortalité France en fonction de l'âge en 1990

Pour identifier la tendance et la volatilité, on analyse le $\text{logit}(q(a, t))$ comme une série temporelle :

$$\begin{aligned} \text{logit}(q(a, t)) &= m_t(a) + \epsilon_t(a) \\ &= m_0(a) + \mu(a)t + \epsilon_t(a) \end{aligned}$$

On pose $\overline{\text{logit}}(a) = \frac{1}{n} \sum_{t=1}^n \text{logit}(q(t, a))$, l'indicateur de la tendance centrale moyenne par âge. La tendance de la série temporelle est donc représentée par le coefficient $\hat{\mu}(a)$ dont la valeur est estimée par la méthode des moindres carrés sur une période de n dates de la manière suivante :

$$\hat{\mu}(a) = \frac{12}{n(n^2 - 1)} \left(\sum_{t=1}^n t \text{logit}(q(a, t)) - \frac{n+1}{2} n \overline{\text{logit}}(a) \right).$$

La variance empirique qui est l'indicateur de dispersion peut se calculer de la manière suivante :

$$\hat{\sigma}^2(a) = \frac{1}{n} \sum_{t=1}^n (\text{logit}(q(a, t)) - \overline{\text{logit}}(a))^2.$$

On va donc chercher à voir si la tendance estimée $\hat{\mu}(a)$ et la volatilité estimée $\hat{\sigma}(a)$, qui sont associées aux paramètres $\mu(a)$ et $\sigma(a)$, sont bien linéaires en l'âge afin de justifier la forme du modèle. Sur une période assez longue ne contenant pas les variations importantes dues aux guerres (1954/2006), on représente la tendance $\hat{\mu}(a)$ et la volatilité $\hat{\sigma}(a)$ de la série temporelle correspondante.

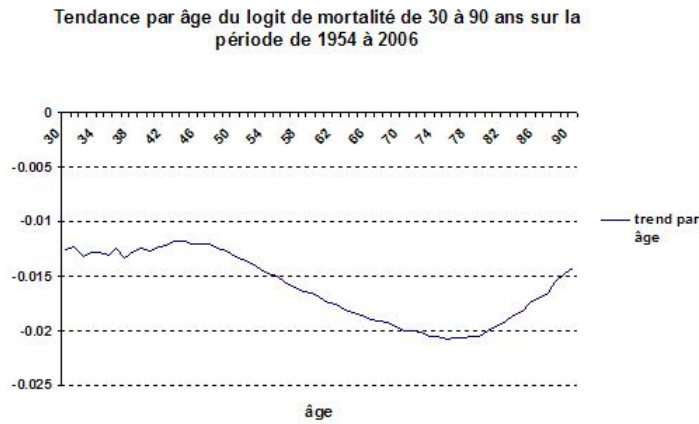


FIGURE 5.8 – Analyse de la pente du logit en fonction de l'âge sur la période 1955/2006

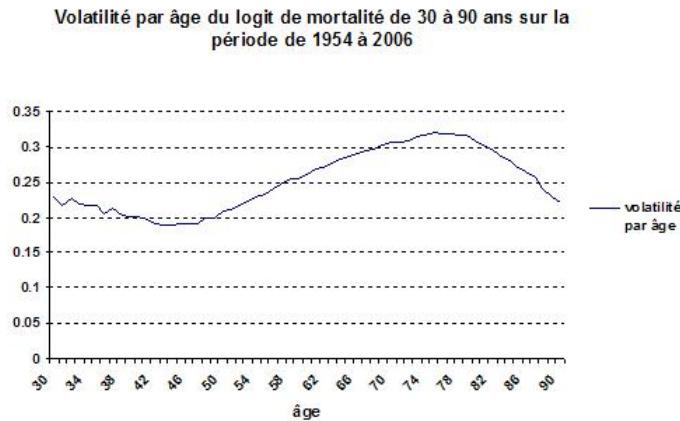


FIGURE 5.9 – Analyse de la volatilité du logit de mortalité en fonction de l'âge sur la période 1955/2006

On vérifie la linéarité en l'âge de la tendance $\mu(a)$ entre 40 ans et 80 ans ainsi que la linéarité en l'âge de la volatilité $\sigma(a)$ entre 40 et 80 ans (voir Figure 5.10 à Figure 5.11).

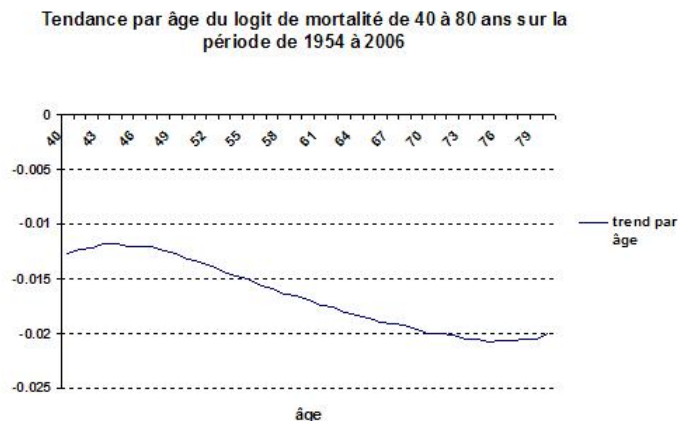


FIGURE 5.10 – Analyse de la pente du logit en fonction de l'âge sur la période 1955/2006

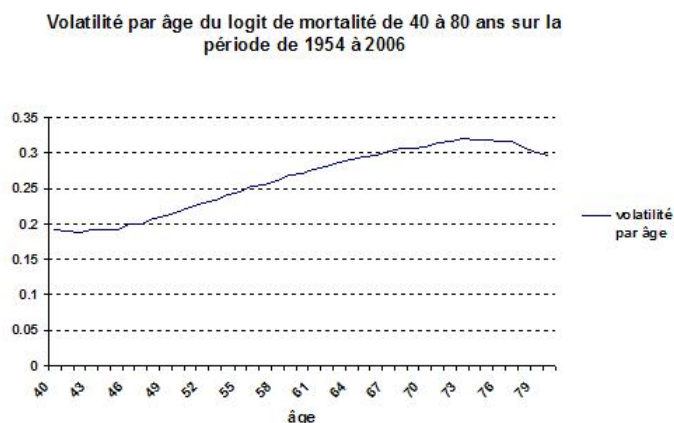


FIGURE 5.11 – Analyse de la volatilité du logit de mortalité en fonction de l'âge sur la période 1955/2006

La linéarité en âge est loin d'être parfaite sur l'ensemble des âges avec certains effets intéressants. On observe que la pente $\mu(a)$ et la volatilité $\sigma(a)$ atteignent respectivement leur minimum et leur maximum vers l'âge de 78 ans. Cela s'explique par le fait que les âges extrêmes ont une amélioration moindre que les âges plus faibles.

On pourrait éventuellement corriger cette non-linéarité en âge de la volatilité de manière à modéliser la mortalité sur une plage d'âge plus étendue. Par exemple, une tendance et une volatilité polynomiales pourraient "fitter" plus précisément les courbes des figures pré-

cédentes sur la tranche d'âge complète 30/90. Cela est laissé pour de futurs travaux. Concernant les simulations, on utilise un procédé de simulation pas à pas qui permet de diminuer l'erreur aux âges faibles :

$$\text{logit}(q(t, a)) = \text{logit}(q(t-1, a)) + [\mu_1 + a\mu_2] + [C^{11}Z_1 + aC^{21}Z_1 + aC^{22}Z_2].$$

L'évolution des facteurs $A_1(t)$ et $A_2(t)$ dans le temps est analysée en étudiant les différentiels notés $\Delta_1(t) = A_1(t) - A_1(t-1)$ et $\Delta_2(t) = A_2(t) - A_2(t-1)$. En effet, ce modèle postule que les différentiels sont gaussiens mais l'historique ne confirme pas forcément ce postulat (voir Figure 5.12 à Figure 5.13). À chaque date t , on extrait les coefficients $A_1(t)$ et $A_2(t)$ par régression linéaire sur les logit de mortalité par âge et on analyse les différentiels de ces coefficients dans le temps.

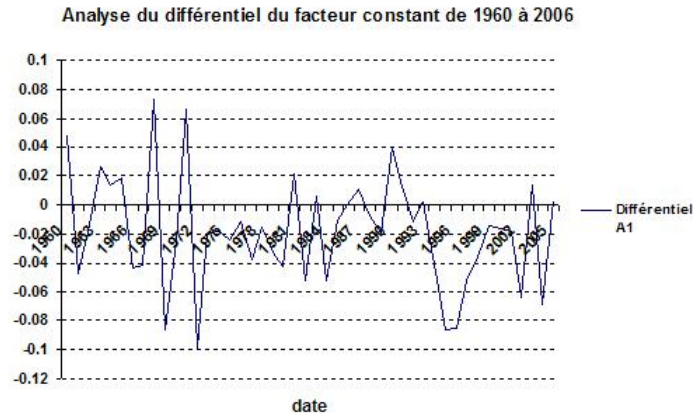


FIGURE 5.12 – Évolution du différentiel du facteur constant

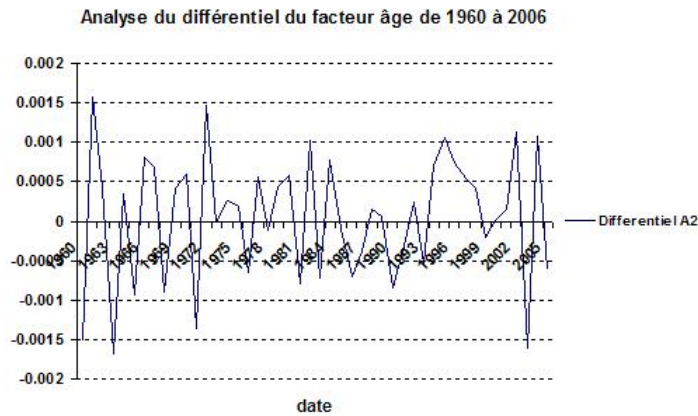


FIGURE 5.13 – Évolution du différentiel du facteur âge

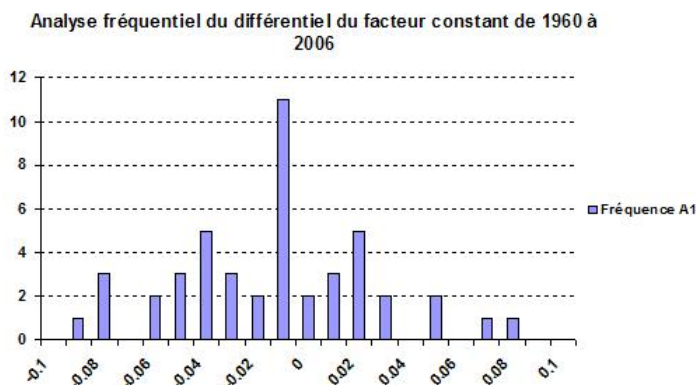


FIGURE 5.14 – Distribution de l'historique du différentiel du facteur constant

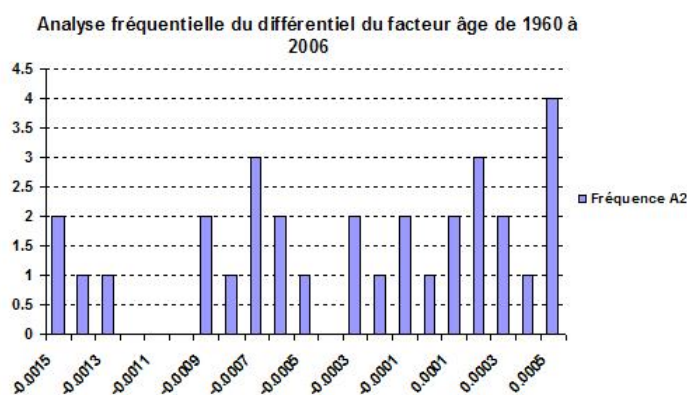


FIGURE 5.15 – Distribution de l'historique du différentiel du facteur âge

Le manque de données ne permet d'identifier précisément la distributions du différentiel des facteurs. Par conséquent, il n'est pas évident de trouver une autre loi qui serait plus adaptée que la loi gaussienne et c'est pour cela qu'on suppose dans un premier temps un modèle gaussien. On peut tout de même noter que le différentiel du facteur A_1 a une distribution assez symétrique et que l'utilisation de la gaussienne n'est pas absurde. Concernant le facteur A_2 , il faudrait sûrement considérer une loi différente de la loi gaussienne.

5.1.2 Calibration du modèle de mortalité et scénarios de projection

Le modèle de mortalité est calibré sur les données historiques de mortalité pour les hommes et les femmes. La calibration s'effectue en considérant pour chaque date, la valeur

des facteurs $A_1(t)$ et $A_2(t)$ ce qui fournit un échantillon de différentiels qui est supposé gaussien. Ensuite, il ne reste qu'à estimer les paramètres μ et Σ en supposant que l'échantillon soit une gaussienne bivariee d'espérance μ et de matrice de variance covariance $\Sigma = C^T C$. La matrice C estimée n'est pas a priori unique (à une rotation près) mais cela ne pose pas de problèmes et on peut prendre la matrice de Cholesky associée, triangulaire inférieur, qui elle est unique.

En calibrant entre 1950 et 2006 pour la tranche d'âge 30/80, on obtient les coefficients suivants :

$$\begin{aligned}\mu^H &= \begin{pmatrix} \mu_1^H \\ \mu_2^H \end{pmatrix} = \begin{pmatrix} -0.01491 \\ -3.113 \times 10^{-5} \end{pmatrix}, \quad \Sigma^H = \begin{pmatrix} \Sigma_{11}^H & \Sigma_{12}^H \\ \Sigma_{21}^H & \Sigma_{22}^H \end{pmatrix} = \begin{pmatrix} 2.488 \times 10^{-3} & -4.001 \times 10^{-5} \\ -4.001 \times 10^{-5} & 8.947 \times 10^{-7} \end{pmatrix}, \\ \mu^F &= \begin{pmatrix} \mu_1^F \\ \mu_2^F \end{pmatrix} = \begin{pmatrix} -0.02208 \\ -2.646 \times 10^{-5} \end{pmatrix}, \quad \Sigma^F = \begin{pmatrix} \Sigma_{11}^F & \Sigma_{12}^F \\ \Sigma_{21}^F & \Sigma_{22}^F \end{pmatrix} = \begin{pmatrix} 2.711 \times 10^{-3} & -4.379 \times 10^{-5} \\ -4.379 \times 10^{-5} & 9.41 \times 10^{-7} \end{pmatrix}.\end{aligned}$$

Remarque 5.1. *Pour avoir une idée des ordres de grandeurs des paramètres, on considère par exemple une femme âgée de 60 ans, $\mu_1^F \geq 10 \times (60\mu_2^F)$. Ainsi, l'amélioration "naturelle" due au facteur A_1 a un impact au moins dix fois plus important sur le logit de mortalité que l'amélioration spécifique à l'âge due au facteur A_2 . Il est donc important de capter l'évolution du facteur $A_1(t)$ dont la distribution est proche d'une distribution gaussienne. Si on suppose les taux de mortalité petits et déterministes, on peut effectuer une approximation :*

$$q(t, a) \sim q(t-1, a)e^{\mu_1 + a\mu_2} \sim q(t-1, a)(1 + \mu_1 + a\mu_2)$$

Cela permet de voir que la tendance μ peut aussi s'interpréter comme une variation relative de taux de mortalité lorsque les taux sont petits.

Les coefficients de tendance μ s'interprètent de la manière suivante : μ_1 est un coefficient d'amélioration de longévité indépendant de l'âge et son signe négatif traduit une baisse du taux de mortalité (ou plutôt du logit qui en est une fonction croissante) dans le temps. μ_2 représente un coefficient d'amélioration de longévité dépendant de l'âge et son signe négatif représente le fait que la baisse de mortalité est plus importante pour les âges élevés. Si on avait calibré le modèle sur la tranche d'âge 80/90, μ_2 aurait été positif dans le cas des hommes en France puisqu'on a vu au graphique Figure 5.8 que la tendance $\hat{m}u$ est une fonction de l'âge qui devient décroissante à partir de 80 ans. Les coefficients de la matrice de variance covariance modélisent les variations de mortalité autour de la tendance.

En ayant calibré le modèle de Cairns, on peut en déduire des projections de mortalité jusqu'en 2047. Par exemple, on représente un scénario d'évolution de la mortalité sur la tranche d'âge 60/80 pour plusieurs dates sélectionnées (voir Figure 5.16). On représente aussi une projection moyenne de l'évolution de la mortalité pour la tranche 60/80 ans (voir Figure 5.17) ainsi que la variance de cette projection (voir Figure 5.18) entre 2007 et 2047.

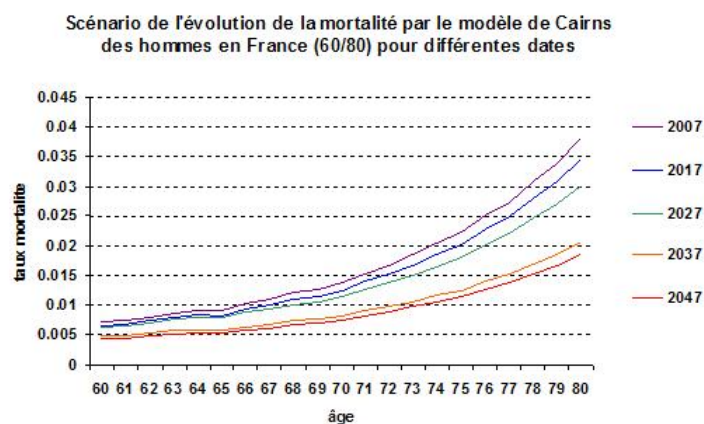


FIGURE 5.16 – Scénarios d'évolution de mortalité entre 2007 et 2047 pour certains âges

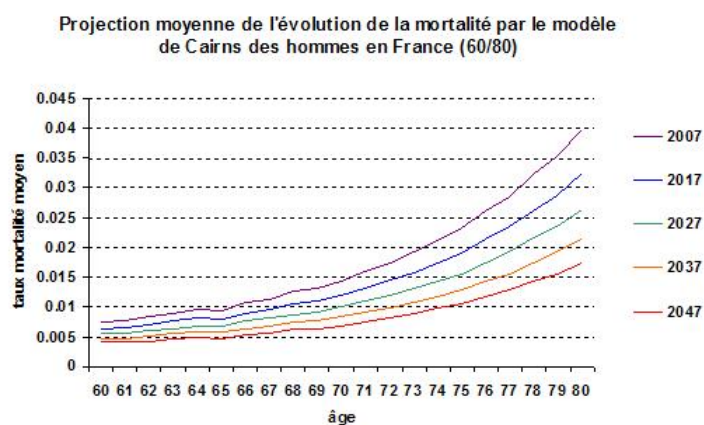


FIGURE 5.17 – Projection Moyenne de mortalité pour la tranche d'âge 60/80 entre 2007 et 2047

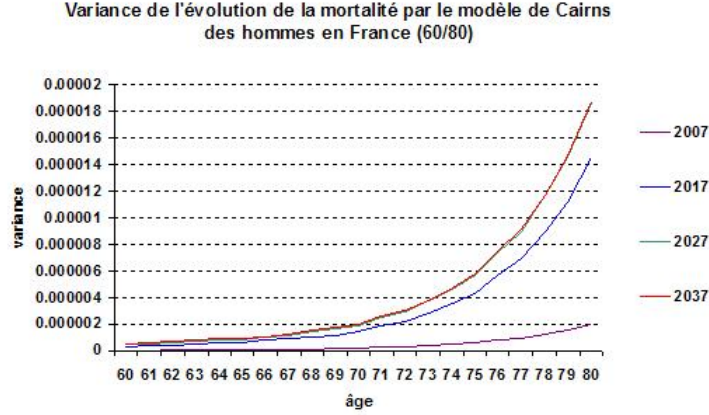


FIGURE 5.18 – Variance de la projection de mortalité pour la tranche d'âge 60/80 entre 2007 et 2047

5.1.3 Backtesting du modèle de Cairns

La procédure de backtesting sur le modèle permet de trouver la période et la tranche d'âge de calibration qui fournissent les meilleurs résultats selon un critère donné. On calibre le modèle en utilisant un ensemble de données, on analyse les projections sur une période dont on connaît la réalisation des taux de mortalité et on compare les résultats du modèle avec les observations réelles. Ainsi, le principe est de changer la plage de données de calibration de manière à obtenir les meilleures projections concernant les tables de mortalité entre 1990 et 2006. Afin de comparer deux courbes de mortalité, on a utilisé un critère qui compare les distances maximales entre les courbes. On s'est principalement intéressé aux âges élevés qui sont au cœur des problèmes de longévité. Les meilleurs résultats ont été trouvés pour la période de calibration 1969/1990 et la tranche d'âge de calibration 30/80.

On peut expliquer intuitivement pourquoi la procédure de backtesting fournit cette période et cette tranche d'âge. La tranche d'âge 30/80 était celle pour laquelle la linéarité en âge était la plus respectée avec une tendance quasi linéaire en l'âge. De plus, pendant la période entre 1969 et 1990, il n'y a pas eu de rupture de pente importante qui pourrait poser des problèmes majeurs. Avec la plage de données issue du backtesting, les paramètres obtenus sont les suivants :

$$\begin{aligned} \mu^H &= \begin{pmatrix} \mu_1^H \\ \mu_2^H \end{pmatrix} = \begin{pmatrix} -0.008538 \\ -1.964 \times 10^{-5} \end{pmatrix}, \quad \Sigma^H = \begin{pmatrix} \Sigma_{11}^H & \Sigma_{12}^H \\ \Sigma_{21}^H & \Sigma_{22}^H \end{pmatrix} = \begin{pmatrix} 1.92 \times 10^{-3} & -3.183 \times 10^{-5} \\ -3.193 \times 10^{-5} & 6.106 \times 10^{-7} \end{pmatrix}. \\ \mu^F &= \begin{pmatrix} \mu_1^F \\ \mu_2^F \end{pmatrix} = \begin{pmatrix} -0.009177 \\ -2.115 \times 10^{-5} \end{pmatrix}, \quad \Sigma^F = \begin{pmatrix} \Sigma_{11}^F & \Sigma_{12}^F \\ \Sigma_{21}^F & \Sigma_{22}^F \end{pmatrix} = \begin{pmatrix} 2.144 \times 10^{-3} & -3.828 \times 10^{-5} \\ -3.828 \times 10^{-5} & 7.576 \times 10^{-7} \end{pmatrix}. \end{aligned}$$

On représente les résultats pour cette calibration sous la forme suivante : un graphique sur la différence entre la projection moyenne et la projection réalisée sur la tranche d'âge 10/70 ans (Figure 5.19) ainsi qu'une représentation plus précise avec la comparaison des taux de mortalité sur plusieurs tranches d'âge : âges bas, âges moyens et âges élevés (Figure 5.20 à Figure 5.22). On remarque d'ailleurs qu'il est difficile de capter la mortalité infantile puisque le modèle de Cairns n'est pas du tout adapté aux âges très faibles pour lesquels la linéarité par rapport à l'âge du logit de mortalité n'est pas vérifiée. (voir Figure 5.7)

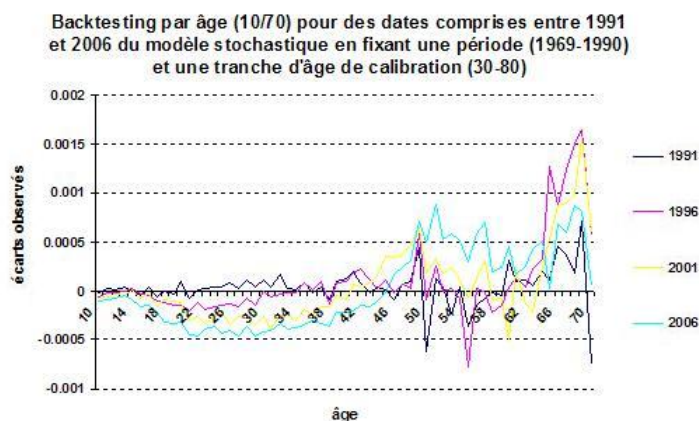


FIGURE 5.19 – Différence entre projection moyenne et projection réalisée (10/70) de 1990 à 2007

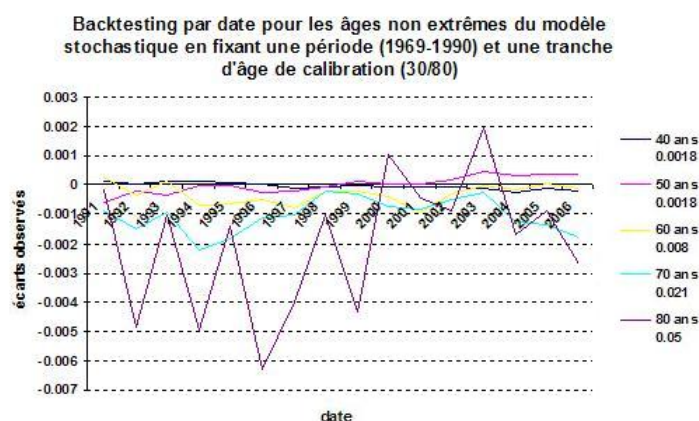


FIGURE 5.20 – Différence entre projection moyenne et projection réalisée (âge moyen) de 1990 à 2007

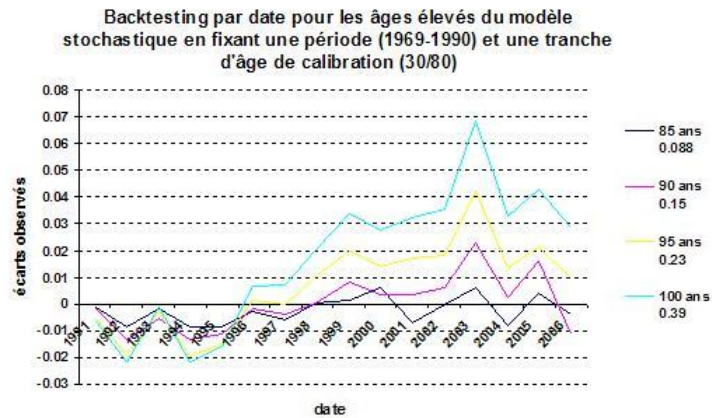


FIGURE 5.21 – Différence entre projection moyenne et projection réalisée (âge élevé) de 1990 à 2007

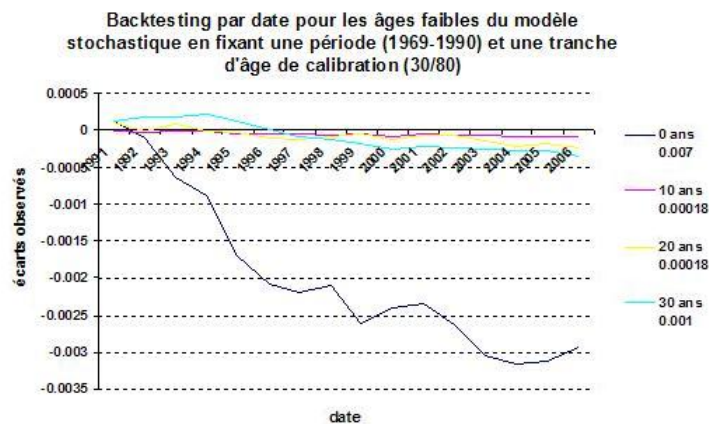


FIGURE 5.22 – Différence entre projection moyenne et projection réalisée (âge bas) de 1990 à 2007

Pour information, on fournit les courbes obtenues pour la même plage de données avec le modèle déterministe associé ou seule la tendance par rapport au temps μ intervient. On se rend compte que la présence d'une volatilité peut permettre de suivre les variations de la courbe de mortalité ce qui fournit des écarts de projections plus faibles.

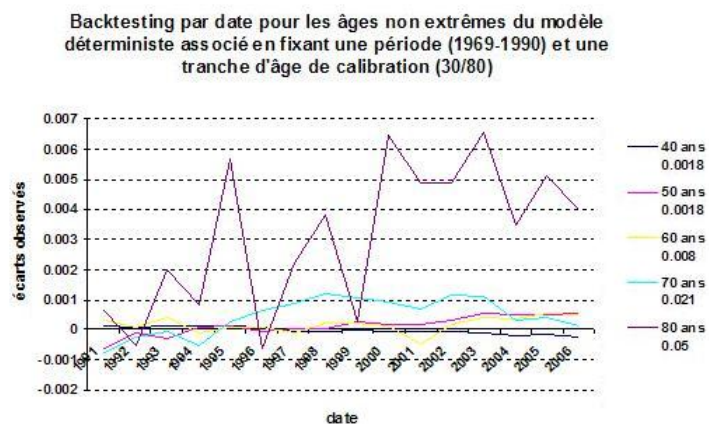


FIGURE 5.23 – Différence entre projection déterministe et projection réalisée (âge moyen) de 1990 à 2007

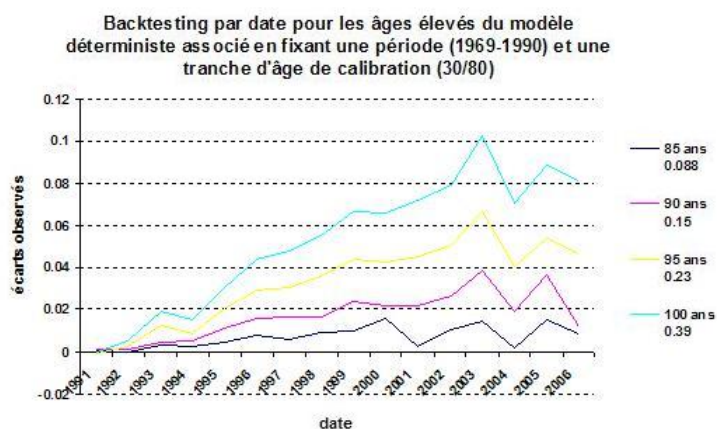


FIGURE 5.24 – Différence entre projection déterministe et projection réalisée (âge élevé) de 1990 à 2007

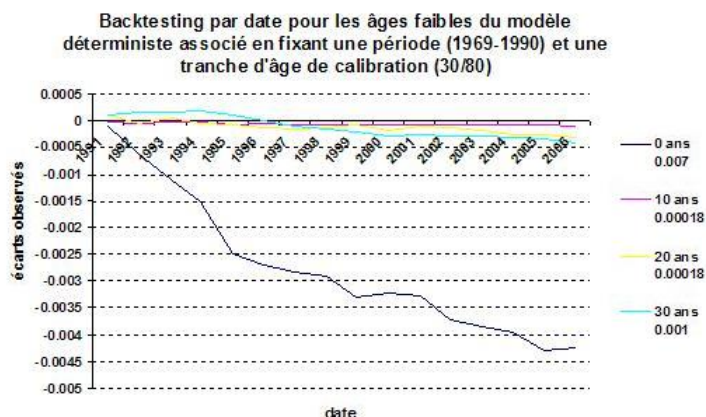


FIGURE 5.25 – Différence entre projection déterministe et projection réalisée (âge bas) de 1990 à 2007

Pour établir une méthode de backtesting efficace, il faudrait un outil mathématique de comparaison de deux courbes plus raffiné dépendant aussi de la tranche d'âge sur laquelle on cherche à être précis. En effet, certains utilisateurs ont besoin d'être le plus précis possible sur la mortalité des âges élevés (comme le détenteur d'un portefeuille de rentes) et d'autres veulent être précis sur l'ensemble de la courbe de mortalité (comme les démographes). Il faudrait éventuellement effectuer une méthode de moindres carrés avec poids pour donner plus d'importance à une plage d'âge spécifique comme celle des âges élevés.

5.1.4 Extension du modèle de Cairns à deux facteurs

Il existe divers extensions de ce modèle de mortalité et nous recensons quelques extensions classiques. Tout d'abord, la première extension naturelle concerne l'ajout de facteurs. En effet, il est possible de tenir compte de la structure de corrélation entre les taux de mortalité de différents âges ce qui permet d'obtenir une dynamique plus réaliste de la courbe de mortalité. Il est aussi possible de modéliser l'effet cohorte ou l'effet générationnel avec une mortalité dépendant de la date de naissance de l'individu.

De plus, il est possible de mieux se caler sur les courbes de mortalité en tenant compte d'un effet quadratique qui permet de tenir compte de la non linéarité en âge (voir Figure 5.7).

Remarque 5.2. Une étude à la Société Générale menée par B. Davalos et al. [1] a montré qu'il pouvait être préférable de modifier le modèle de Cairns en prenant un facteur proportionnel à une puissance de l'âge et non pas linéaire en l'âge. Cela permettrait d'avoir un ajustement plus précis du modèle sur les courbes de mortalité, en ayant un levier sur la convexité par rapport à l'âge de la courbe. Pour cela, on définit $\alpha(a) = c_2 a^2 + c_1 a + c_0$ et

on écrit le modèle suivant :

$$\text{logit}(q(a, t)) = \text{logit}(q(a, t - 1)) + \mu_1 + \sigma_1 Z_1(t) + (\mu_2 + \sigma_2 Z_2(t)) a^{\alpha(a)}.$$

5.1.5 Extension à un modèle prenant en compte les catastrophes

En analysant les courbes de mortalité, on observe des variations naturelles autour d'une tendance avec certaines variations exceptionnelles expliquées par des événements majeurs : pandémies, guerres, catastrophe naturelles. Il apparaît important de prendre ces phénomènes en compte dans la modélisation de la mortalité. Cette surmortalité exceptionnelle peut être modélisée par l'occurrence d'un événement apparaissant à une date aléatoire et d'intensité aléatoire. Il existe en effet des catastrophes majeures comme la pandémie de grippe espagnole qui ont eu un impact largement plus important que d'autres pandémies dans le monde.

La fréquence historique de ces événements au cours des derniers siècles ainsi que leur intensité peuvent être analysées. Certains experts de l'OMS estiment qu'il y a eu 31 événements de mortalité catastrophique sur les 420 dernières années soit en moyenne de 7.38 événements par siècle. En ce qui concerne la sévérité, il y a eu 3 pandémies majeurs durant le 20^e siècle avec la grippe espagnole (1918 et 1919), asiatique (1957) et de Hong Kong (1968). En moyenne, la sévérité est estimée à 0.3% de surmortalité.

La modélisation classique est de modéliser la surmortalité due à ces événements sous la forme d'un produit d'une composante fréquence par une composante intensité. Une méthode de prise en compte de la surmortalité due aux catastrophes naturelles est décrite dans un travail de L. Devineau [13]. La surmortalité est modélisée comme suit :

$$\begin{aligned} q(a, t) &= q^{\text{central}}(a, t) + B_t(p)E_t(\lambda) \\ \text{logit}(q^{\text{central}}(a, t)) &= A_1(t) + aA_2(t), \end{aligned}$$

dans laquelle la variable B_t suit une loi de Bernoulli à valeur dans 0, 1 tirée à l'instant t dont la probabilité $p = \mathbb{P}(B_t = 1)$ correspond à la probabilité d'occurrence de la catastrophe. La variable E_t suit une loi exponentielle dont le paramètre λ correspond à l'inverse de l'intensité moyenne de la catastrophe (la moyenne d'une loi exponentielle est égale à l'inverse de son paramètre).

Remarque 5.3. *L'inconvénient de ce modèle est qu'il autorise en théorie d'avoir des taux de mortalité supérieurs à 1 puisque la surmortalité est définie sur le taux de mortalité lui même. Une alternative consisterait à modéliser une surmortalité à l'échelle du logit de mortalité sous la forme :*

$$\text{logit}(q(a, t)) = A_1(t) + aA_2(t) + B_t(p)E_t(\lambda)$$

En pratique, le premier modèle de Devineau maintient un niveau de mortalité compris strictement entre 0 et 1 et nous l'avons utilisé dans la simulation.

Afin de calibrer les paramètres de catastrophes naturelles (p et λ), on analyse l'ensemble des catastrophes survenues au cours des derniers siècles avec leur fréquence et leur intensité. Dans nos simulations, la probabilité p utilisée est $p = \frac{7.38}{100} = 0.0738$ et l'intensité moyenne des catastrophes est $\lambda = \frac{3}{1000} = 0.003$.

Lors de la procédure de calibration de la volatilité des facteurs du modèle, les données concernant les périodes de catastrophes naturelles n'ont pas été prises en compte (car cela biaiserait artificiellement la volatilité des périodes normales). En fait, en supposant la volatilité constante dans le temps, ce qui se justifie sur les données, on doit faire face au problème de la rupture de tendance qui contraint a priori à prendre un nombre très limité d'observations sur une période où la tendance est constante. On peut en fait calibrer la volatilité sur l'ensemble des données (hors catastrophes) en déterminant les instants où la tendance change et en rassemblant les données des périodes pendant lesquelles la tendance est quasi constante. En considérant la tendance sur chacune des périodes, on obtient une estimation globale de la volatilité des facteurs de mortalité.

Remarque 5.4. *Une intuition en analysant les tables de mortalité serait qu'une catastrophe naturelle modifie la tendance de la mortalité. Cela est une conjecture qui est seulement intuitive à la vue des tables de mortalité. Si cela se révélait pertinent, on pourrait généraliser nos projections tendanciellles à des projections où la tendance varie à certaines dates qui seront a priori des dates de catastrophes naturelles. Il faut donc tester statistiquement cette conjoncture qui n'est peut être pas totalement pertinente...*

On considère un scénario de projection de la mortalité jusqu'en 2050 avec d'éventuelles pandémies et on le compare à un autre scénario de projection sans pandémie (voir Figure 5.26 et Figure 5.27). Ces scénarios révèlent des pics de surmortalité à différentes dates dont l'intensité est la même pour tous les âges.

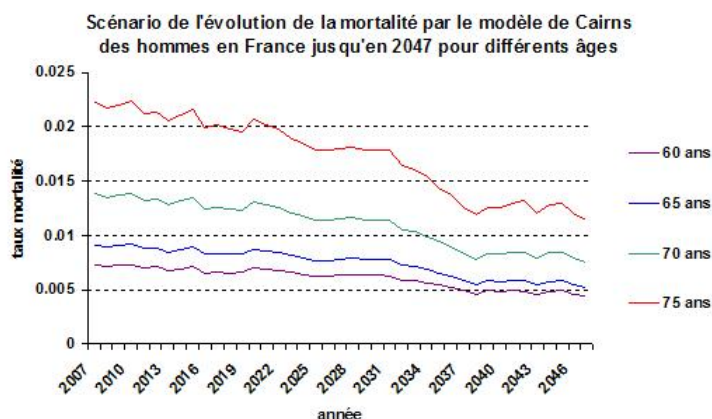


FIGURE 5.26 – Scénarios d'évolution de mortalité entre 2007 et 2047 pour certains âges

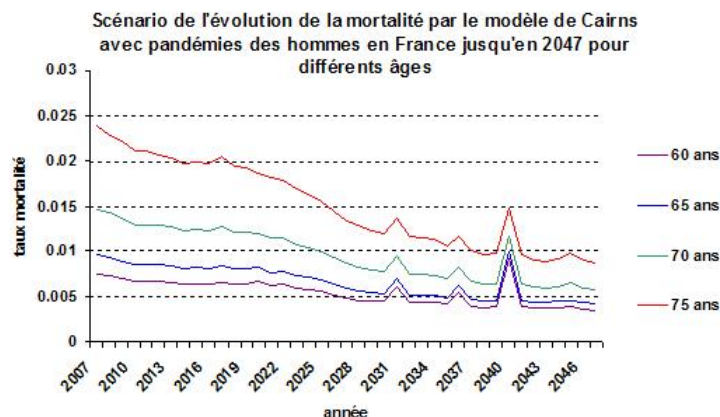


FIGURE 5.27 – Scénarios de mortalité avec prise en compte catastrophe naturelle jusqu'en 2047.JPG

Les sauts visibles sur les courbes de mortalité durant certaines périodes (Figure 5.1 à Figure 5.4) résultent de catastrophes naturelles ou humanitaires comme la grippe espagnole ou les deux guerres mondiales (première et deuxième guerre mondiale en 1914/1920 et 1939/1947). Les pandémies étant de nature très différente des guerres, une modélisation de la surmortalité pour ces deux types d'événements est difficile à étudier statistiquement. Les pandémies et les guerres n'affectent pas tous les âges de la même manière (les sauts sont d'intensités différentes selon les âges) et une modélisation plus fine serait à faire puisqu'on a considéré que l'impact d'une pandémie était le même sur l'ensemble des âges.

Il est possible de trouver une probabilité d'occurrence, une intensité moyenne et éventuellement une durée de la catastrophe (plusieurs années que ce soit pour les guerres ou pandémies). Il faudrait affiner la modélisation en prenant en compte les répercussions temporelles des catastrophes dans la mesure où la surmortalité se répartit sur plusieurs années. En effet, il y a un effet classique de cycle (de type U ou W) comme par exemple le cycle d'une forte surmortalité suivie d'une faible surmortalité et à nouveau suivie d'une forte surmortalité.

5.2 Étude des facteurs explicatifs de la mortalité

Dans cette partie, nous cherchons à identifier les facteurs explicatifs de la mortalité. De nombreuses études permettent de déduire des facteurs prépondérants de mortalité : ces facteurs seront à la fois des caractéristiques individuelles et des caractéristiques globales concernant l'environnement (social ou économique) dans lequel l'individu évolue. Une étude particulière menée par F. Jusot [17] fournit des résultats intéressants sur l'influence du

niveau socioéconomique sur la mortalité et nous y ferons souvent référence.

5.2.1 Sexe de l'individu

Il faut tenir compte de la différence de sexe pour plusieurs raisons. Tout d'abord, il existe des différences significatives entre la mortalité des hommes et des femmes (voir Figure 5.28) et ce trait va être considéré comme un trait caractéristique majeur. En 2005, l'espérance de vie en France à la naissance était estimée à 76.8 ans pour les hommes contre 83.8 ans pour les femmes. Ces écarts se maintiennent aux âges élevés puisqu'à 60 ans, l'espérance de vie résiduelle des hommes était de 21.4 années contre 26.4 années pour les femmes.

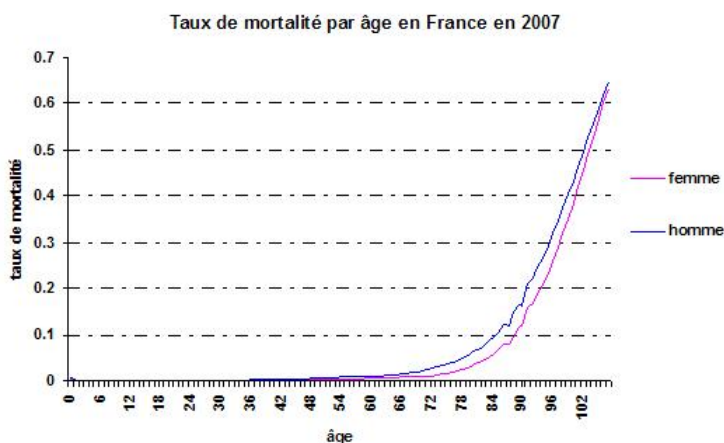


FIGURE 5.28 – Taux de mortalité par âge pour la population française en 2006

On pourrait penser que cet écart de mortalité homme/femme est expliqué uniquement par des différences de conditions de vie et de travail. L'analyse de la mortalité infantile permet de réfuter cette théorie. En effet, la mortalité infantile en France révèle aussi une différence significative de mortalité entre les deux sexes (voir Figure 5.29). Ce phénomène pourrait s'expliquer par le fait que les femmes seraient génétiquement favorisées par rapport aux hommes. Cette remarque permet de considérer le sexe comme un trait explicatif majeur de mortalité.

Remarque 5.5. *On a traité le cas d'un pays comme la France : cette remarque ne sera plus vraie pour un pays comme la Chine pour lequel il existe des inégalités de mortalité liées aux conditions de vie à la naissance différentes pour les filles et les garçons.*

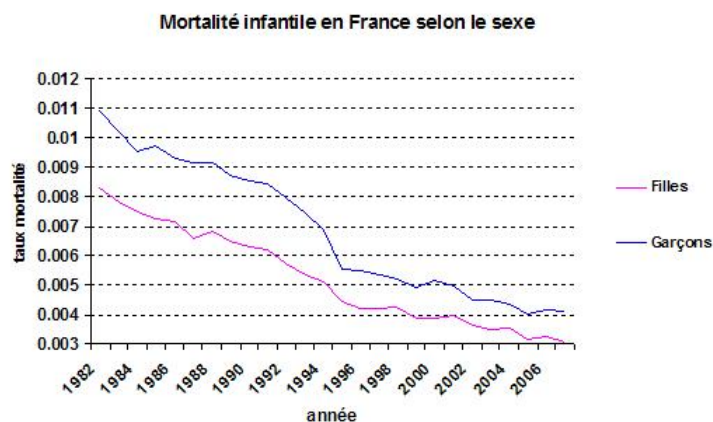


FIGURE 5.29 – Taux de mortalité infantile par sexe en France

5.2.2 Statut matrimonial et vie familiale

Des études récentes montrent que la vie familiale joue un rôle protecteur face au décès [21]. La vie familiale peut être analysée par le statut matrimonial ainsi que le nombre d'enfants du foyer. En effet, la mortalité des femmes n'ayant pas connu de vie de couple et n'ayant pas de descendance est plus élevée que celle des femmes ayant vécu une vie de couple ou qui ont eu des enfants. À l'âge de 45 ans, une femme célibataire a une surmortalité de 26% par rapport à l'ensemble des femmes et une femme qui n'a pas eu d'enfants a une surmortalité de 14%.

Ces études soulignent que le mariage a un effet protecteur ainsi qu'une surmortalité pour la classe des célibataires, des divorcés et des veufs. Le mariage "protège" à la fois les hommes et les femmes avec des ordres de grandeur comparables. La surmortalité des célibataires peut provenir soit d'un effet de sélection par la santé sur le marché du mariage (à l'instar de la dépendance) soit d'un effet néfaste du célibat par rapport au mariage. Le caractère protecteur de la vie familiale est-il intrinsèque ou résulte-t-il d'une situation économique plus favorable ?

Il a été démontré que le statut matrimonial est corrélé au risque de décès et qu'il explique encore significativement la probabilité de décès après un contrôle du niveau de revenu [17]. À autres caractéristiques identiques (âge, taille du ménage, statut d'occupation du logement, diplôme), les risques annuels de décès sont plus faibles pour les hommes et femmes mariées. Ainsi, la corrélation entre mariage et longévité a un impact direct et n'est pas due à une corrélation entre mariage et statut économique comme on aurait pu le penser.

Concernant le nombre d'enfants, on sait que le fait d'avoir des enfants diminue le niveau

mortalité de l'individu [21]. L'effet est nuancé et il existe de grandes variations selon le nombre d'enfants mis au monde. Les femmes ayant un enfant unique ont un risque de décès de 3% inférieur et le risque de décès diminue à 10% pour 3 enfants qui est la situation familiale où la sousmortalité est maximale. Au delà de 4 enfants, cet avantage s'amenuise et le fait d'avoir plus de 5 enfants devient un facteur de surmortalité.

5.2.3 Catégorie socioprofessionnelle

La catégorie socioprofessionnelle a un impact sur le niveau de mortalité aux âges actifs. Cet impact est moins marqué chez les femmes que chez les hommes : la structure professionnelle a davantage d'effets sur la mortalité des hommes que sur celle des femmes [31]. Chez les hommes, il existe des différences importantes de mortalité entre deux catégories sociales différentes. Par exemple, en ce qui concerne l'espérance de vie des français de 35 ans, il y a une différence de plus de 6 ans entre les cadres et les ouvriers hommes alors qu'elle n'est que de 3.5 années chez les femmes. Il faudra par conséquent trouver une autre variable pour étudier la situation sociale de la femme comme le niveau de diplôme.

Aux âges élevés, les différences de mortalité diminuent très fortement et deviennent rapidement négligeables [30]. En France, l'impact des catégories socioprofessionnelles est plus important pour les âges inférieurs à 80 ans. Cela résulte sûrement du fait que le système de santé en France est basé sur un concept de solidarité où les soins sont souvent pris en charge par l'état. D'ailleurs, cette observation n'est pas valable aux États Unis où les écarts restent importants aux âges élevés.

La catégorie socioprofessionnelle est assez difficile à analyser dans la mesure où elle peut changer de nombreuses fois au cours de la vie d'un individu et qu'elle n'est donc pas stable dans le temps. En effet, même si on connaît la catégorie socioprofessionnelle du défunt, si ce dernier a changé de catégorie sociale au cours de sa vie, il est difficile de lui associer une catégorie sociale unique. Par exemple, il faut trouver des moyens de déterminer la catégorie socioprofessionnelle d'un homme qui était un employé la majeure partie de sa vie et qui est décédé en tant que cadre. La classification des catégories socioprofessionnelles ayant aussi évolué ces dernières années, il est difficile d'avoir une classification de référence.

5.2.4 Niveau d'éducation

Le niveau de diplôme joue un rôle important sur le niveau de mortalité, que ce soit chez l'homme ou la femme avec une prédominance chez la femme. Contrairement à l'influence de la catégorie sociale qui n'est pas beaucoup marquée chez la femme, le niveau de diplôme a un impact significatif chez la femme.

De plus, la variable du niveau de diplôme est stable dans le temps puisqu'elle est généralement figée à partir de 30 ans. Ainsi, il sera plus pertinent d'analyser les différentiels sociaux de mortalité au travers du niveau d'éducation plutôt qu'au travers de la catégorie sociale.

Aux âges élevés, les écarts diminuent fortement et s'estompent petit à petit mais ils restent tout de même significatifs [30]. En 2002, les femmes de 86 ans ayant un diplôme supérieur ou égal au baccalauréat vivaient encore 7.1 années contre 3.1 pour les femmes sans diplôme. Il en est de même pour les hommes pour lesquels les diplômés peuvent aussi espérer à cet âge vivre encore 20% plus longtemps que les non diplômés.

5.2.5 Niveau de revenu

Dans cette partie, nous étudions l'impact du niveau de revenu sur la mortalité. Cette étude est assez complexe et il faut définir le concept économique du revenu, comprendre l'impact du revenu sur le niveau de santé et voir ce que la variable du revenu peut apporter par rapport à la variable de catégorie socioprofessionnelle.

Il y a deux points de vue économiques sur la pauvreté : le premier point de vue est celui du concept de pauvreté absolue où l'on souligne l'impact du niveau de revenu sur la santé (non accès aux soins, logements précaires). Le deuxième point de vue est celui du revenu absolu selon lequel il n'y a plus seulement opposition des pauvres et non pauvres (mise en évidence par les catégories socioprofessionnelles) mais il y a aussi une prise en compte des revenus disponibles pour chaque foyer ce qui explique le gradient économique de santé. Il existe un effet du rang de l'individu dans la société sur sa santé. On considérera dans la suite les revenus en unité de consommation (par foyer) qui sont introduits en quintiles afin de différencier des tranches de revenus et de tenir compte de toute la distribution des revenus.

Impact du revenu sur la santé

Les modèles démographiques de mortalité considèrent l'évolution de la santé d'un individu i . La mort d'un individu est un événement qui se produit lorsque la santé de l'individu H_i passe en dessous d'un certain seuil H_{min} . Il existe des modélisations démographiques qui prennent en compte la santé et le revenu ainsi que leur corrélation. On peut citer un modèle en exemple où H correspond à la santé et Y au revenu.

$$\begin{cases} H_{it} &= a_0 + a_1 Y_{it} + a_2 X_{it}^h + a_3 Z_{it}^h + u_{it} \\ Y_{it} &= b_0 + b_1 H_{it} + b_2 X_{it}^y + b_3 Z_{it}^y + v_{it}. \end{cases}$$

Ce modèle tient compte du fait que la santé est une fonction du revenu mais aussi que la santé a elle même un effet inverse sur le travail. La santé détermine le temps de travail et

le travail devient plus pénible lorsque la personne a des problèmes de santé. La santé est souvent considérée comme une fonction croissante du revenu jusqu'à un certain seuil de revenu (au delà d'un certain niveau, le revenu n'a plus d'impact sur la santé). Finalement, il existe un lien entre la santé et les revenus : les inégalités sociales de santé sont expliquées par le statut socioéconomique de l'individu au travers du revenu et de la profession.

Il a été prouvé que le revenu a un impact sur l'ensemble de la population à tous les âges : les pauvres révèlent une surmortalité mais on souligne aussi une sous mortalité chez les riches. En France, la corrélation revenu/décès est forte mais celle-ci diminue lorsque l'âge des individus augmente.

Différences entre revenus et catégories socioprofessionnelles

On pourrait penser que la CSP donne déjà de l'information sur le niveau de revenu et que les variables sont redondantes. Cependant, la CSP ne fournit pas une relation directe entre le revenu et la mortalité : en effet, au sein même d'une même CSP, la variance de revenu est deux fois supérieure en comparaison de la variance de revenu qu'il existe entre deux CSP différentes. Ainsi, le lien entre les revenus et la mortalité est différente de celle qu'il existe entre la CSP et la mortalité. L'élasticité de la probabilité de décès au revenu est de l'ordre de -0.4 , cet effet étant plus marqué avant 65 ans (-0.5) qu'après 65 ans (-0.3) [17]. Il est démontré que le revenu est un facteur économique explicatif. En effet, l'impact des revenus, bien que réduit (l'élasticité est de l'ordre de -0.3), persiste après contrôle des catégories socioprofessionnelles. D'ailleurs, la probabilité de décès diminue avec le niveau de revenu au sein de chaque profession, le risque des cadres restant inférieur à celui des autres CSP.

Une étude économétrique montre que le revenu a un effet positif sur la santé et permet de diminuer les risques de décès [17]. Il y a un effet distinct de la profession et de la CSP : en fait, le revenu par unité de consommation ne diminue que faiblement le caractère explicatif de la CSP et il y a presque indépendance entre ces deux caractéristiques.

La corrélation revenu/décès est forte mais diminue aux âges élevés en France. Ce constat n'est pas fait aux États Unis où l'accès à Medicaid est différent de celui à la sécurité sociale en France.

Le décès du premier quintile de revenu est aussi bien plus grand que le décès des autres quintiles ce qui explique le caractère très néfaste de la pauvreté. Sans prendre en compte la profession, on observe 2.5 fois plus de décès du dernier quintile (revenus les plus bas) par rapport à l'avant dernier quintile. En tenant compte de la profession, on observe 2 fois plus de décès entre ces deux quintiles. Enfin, on souligne une sous mortalité des revenus les plus

élevés (premier quintile) par rapport à celle du deuxième quintile. Ainsi, non seulement la pauvreté augmente le niveau de mortalité, mais la richesse est protectrice même par rapport à un niveau de revenu relativement aisé.

Effet des inégalités de revenus

Une première analyse montre que la fonction de survie est croissante du revenu pour tous les âges. S.H. Preston [24] prouve, dès 1975, que l'espérance de vie n'augmente pas continûment avec le revenu par tête. Alors que dans les pays en voie de développement, la richesse moyenne explique largement l'état de santé moyen, cette relation n'est plus vraie dans les pays riches.

Quelques années plus tard, G.B. Rodgers [32] a cherché à expliquer la mortalité moyenne par le niveau des inégalités de revenu sous l'hypothèse selon laquelle les inégalités de revenu correspondent aussi aux inégalités de santé. En 1992, R.G. Wilkinson [36] a démontré que "la santé est fortement affectée par l'étendue des différences économiques et sociales au sein d'une population". En effet, son étude révèle une corrélation entre l'espérance de vie et l'indice de Gini (indice représentant un niveau d'inégalités au sein d'une population). Le problème de cet indice est qu'il tient compte de 11 pays mais qui ont des politiques de santé et un contexte socioculturel différents. Il existe trois hypothèses concernant l'effet des inégalités de revenus sur la santé que nous allons détailler et analyser.

La première hypothèse est l'hypothèse de revenu absolu qui analyse la relation existant entre revenu et la santé. En effet, il existe une relation concave entre le revenu et la santé (Rodgers) ce qui s'explique qualitativement : on sait que la santé est une fonction croissante du revenu, mais cette fonction est aussi à rendements décroissants ce qui souligne l'impact des inégalités de revenu. En effet, le caractère concave souligne le fait qu'une inégalité sur les revenus induit une baisse de l'état de santé moyen. L'amélioration de l'état de santé d'un riche, par un transfert de revenu, est inférieure à une détérioration d'un état de santé d'un pauvre due à la perte d'un revenu de même valeur. Wilkinson rejette cette hypothèse dans la mesure où une inégalité convexe entre la mortalité et le revenu apparaît insuffisante car elle ne permet pas d'expliquer la corrélation existante entre la mortalité et les inégalités de revenu.

Wilkinson propose une deuxième hypothèse selon laquelle l'état de santé individuel n'est pas expliqué par le niveau de revenu absolu mais par la position relative de l'individu au sein de la société. Cette hypothèse traduit des mécanismes psychosociaux, et les différences d'état de santé ne seraient pas expliquées uniquement par le niveau de revenu des individus mais aussi par l'écart entre leur revenu et le revenu d'un groupe de référence. Cet effet s'explique par la cohésion sociale sur la santé ou par l'influence des comparaisons sociales sur

la santé. En effet, on peut expliquer cela en considérant qu'un niveau moindre de cohésion sociale ou de capital social a un impact négatif sur la santé des individus. Il existe d'ailleurs un indicateur du capital social donné par R.D. Putnam [25] en 1995. La corrélation peut d'ailleurs être expliquée par des caractéristiques géographiques comme les décisions locales.

Enfin, la dernière hypothèse envisage un impact indirect des inégalités. La corrélation pourrait refléter, non pas un effet direct des inégalités, mais un impact des caractéristiques géographiques, elles mêmes corrélées aux inégalités, telles que les décisions politiques locales [10]. Par exemple, on peut constater que les États américains ayant une distribution plutôt inégalitaire des revenus ont un taux de mortalité plus élevé, un taux de chômage plus élevé, un taux de criminalité en progression, un nombre de personnes non assurées élevé ainsi qu'un niveau d'éducation faible. Il faut analyser ces différences de mortalité expliquées seulement par des caractéristiques individuelles mais aussi celles dues aux caractéristiques de l'environnement social dans lequel évolue l'individu (au travers du niveau des inégalités par exemple).

La France est un pays très adapté à l'étude des inégalités sociales. L'étude de F. Jusot [17] a été de voir si la corrélation entre les inégalités de revenu et l'espérance de vie est un artefact statistique résultant de la relation concave entre santé et revenu ou si les inégalités de revenu constituent l'un des déterminants de la santé individuelle.

L'étude de F. Jusot révèle que la première hypothèse n'est pas validée dans la mesure où la relation de concavité n'est plus respectée : elle a démontré l'effet attendu d'une surmortalité liée à la pauvreté ainsi qu'un effet significatif moins attendu d'une sous mortalité des personnes les plus favorisées (effet protecteur de la richesse). Le risque de décès des 20% des plus pauvres est 2.5 fois plus élevé que le risque de décès des plus riches (le rapport se ramène à 2 fois plus élevé après contrôle des CSP).

La deuxième hypothèse n'est pas validée dans le cas de la France. En effet, l'étude sur la criminalité, qui est un impact du capital social, révèle que cette dernière n'est pas un facteur explicatif de risque de décès. Le taux de criminalité étant théoriquement corrélé au niveau de cohésion social, l'effet du capital social n'est pas directement validé. Ce résultat réfute la deuxième hypothèse selon laquelle l'impact des inégalités sociales serait du à un changement de cohésion social.

F. Jusot a démontré que c'était la troisième hypothèse qui était validée dans le cas de la France avec un impact indirect des inégalités de revenus. Il faut tenir compte des caractéristiques régionales telles que l'offre de soin et le taux de chômage qui reflètent certaines décisions politiques.

Influence de la région d'habitat

À présent, on s'intéresse à la variabilité de décès interrégionale. En l'absence des caractéristiques individuelles, on obtient une variance interrégionale de mortalité non nulle. Cette dernière reste non nulle lorsqu'on introduit les caractéristiques individuelles. Un moyen de la faire diminuer est d'introduire des variables régionales comme les inégalités économiques, le taux de chômage et l'offre de soin. Le risque de décès apparaît être une fonction croissante des inégalités économiques et une fonction décroissante de la densité médicale. Le problème de l'offre de soin est qu'elle apporte une confusion dans la mesure où les régions les plus inégalitaires sont celles qui ont la densité médicale la plus élevée (Île de France, PACA). Il faut considérer des groupes de régions qui appartiennent à une même tranche d'inégalité.

La corrélation entre le soin et le décès apparaît être élevée et celle entre le chômage et le décès l'est moins puisqu'il y a déjà une prise en compte du statut à travers les variables individuelles (avec des professions plus touchées que d'autres par le chômage). La prise en compte de ces deux variables d'inégalités permet de réduire fortement la variance interrégionale. Les résultats diffèrent selon les indicateurs d'inégalité : la mesure d'entropie de Theil et l'indice d'Atkinson donnent de bons résultats. Enfin, on précise que les inégalités sont généralement plus importantes parmi les individus les personnes ayant un revenu inférieur à un revenu médian.

5.2.6 Étude du déterminisme social et génétique

Une étude menée par l'IRDES avec M. Devaux et al. [8] concerne les inégalités de chance en santé avec l'influence de la profession et de l'état de santé des parents. En effet, ils s'interrogent sur le déterminisme social et familial dans la mesure où la santé à l'âge adulte dépend du milieu social d'origine.

Il existe trois hypothèses sur l'influence du patrimoine social et génétique. La première hypothèse suggère un effet direct des conditions de vie dans l'enfance sur la santé à l'âge adulte. La seconde suppose plutôt un effet indirect de l'environnement familial sur le futur statut socioéconomique. La dernière correspond à la transmission de l'état de santé entre les générations fondée sur un patrimoine génétique commun avec une reproduction des comportements à risque, des recours aux soins et de prévention.

Cette étude considère le test de ces trois hypothèses sur des données françaises et les résultats suggèrent que l'état de santé à l'âge adulte est directement influencé par la profession de la mère alors que la profession du père a une influence indirecte passant par la détermination du statut socioéconomique de l'enfant. De plus, sans annuler l'influence du milieu social, l'état de santé des deux parents a un effet direct sur la santé de l'enfant à l'âge adulte. Le niveau d'éducation n'est pas non plus à négliger dans la mesure où il permet de

réduire nettement les inégalités sociales de santé à l'âge adulte.

5.2.7 Dépendance des individus et prévention

L'entrée en dépendance, qui concerne généralement les personnes âgées, est un enjeu démographique majeur. Ce phénomène est typiquement français, le mot "dependency" ou "loss of autonomy" n'existant pas en anglais.

Il y a officiellement 1 078 000 personnes dépendantes en France et les scénarios de projections envisagent une évolution lente de 1% par an. En effet, l'augmentation de l'espérance de vie sans incapacité fait baisser le taux de croissance de la proportion des personnes dépendantes au sein de la population. Cependant, le nombre des personnes âgées de plus de 75 ans s'accroît sensiblement ce qui entraîne tout de même une augmentation du nombre de personnes dépendantes. D'ici 2040, la population des plus de 75 ans devrait dépasser 10 millions de personnes et la population des plus de 80 ans constituerait environ 7 millions de personnes. Ces projections envisagent pour 2040 qu'il y aurait 1 230 000 personnes âgées dépendantes. La population dépendante serait plus âgée qu'aujourd'hui, avec une augmentation moyenne de 4 ans pour les hommes, 5 ans pour les femmes.

Un individu devient dépendant souvent à la suite d'une maladie s'exprimant chez ce dernier et l'empêchant d'être autonome. En France, des moyens sont mis en place pour aider ces personnes. Après les accidents cardiovasculaires et les cancers, les maladies émergentes pour la nouvelle génération des personnes âgées concernent la démence sénile comme la maladie d'Alzheimer qui concerne 856 000 patients. Dans les formes modérément sévères à sévères de la maladie, la proportion des patients dépendants est de 90% soit plus de 300 000 personnes.

Les maladies de démence sénile sont la cause majeure de l'entrée en dépendance. Les autres causes de dépendance étant les maladies cardiovasculaires, les cancers, l'arthrose et l'ostéoporose, les troubles sensoriels, les dépressions et les troubles nutritionnels. Les prévisions concernant ces maladies sont très alarmantes et selon une hypothèse de prévalence constante, d'ici 2040, 2 150 000 personnes seront atteintes de la maladie d'Alzheimer.

On peut considérer que la dépendance reflète largement le capital santé de l'individu. En effet, les statistiques (INSEE) révèlent que lorsqu'une personne devient dépendante, cette dernière voit inévitablement son espérance de vie se réduire à environ 4 ans, ceci presque indépendamment de l'âge d'entrée en dépendance. Ces statistiques sont obtenues en prenant toutes les formes de dépendance comme celles liées à la maladie d'Alzheimer pour laquelle certaines personnes restent dépendantes pendant de nombreuses années. Ainsi, la longévité des individus est fortement corrélée au niveau de dépendance des individus et il faudra en tenir compte dans un modèle de mortalité.

Le problème de la dépendance est aussi un enjeu économique dans la mesure où une année de dépendance a un coût quatre fois plus élevé qu'une année de retraite classique, et il apparaît très important de prendre des mesures de prévention pour éviter que ce phénomène se généralise.

G. Vallancien, professeur de Médecine reconnu, a proposé quatre moyens de prévention de la dépendance. D'abord, il ne faut pas infantiliser les personnes âgées pensant que cette "dégradation les entraîne dans une déchéance irrépressible". Ensuite, il faut arrêter de médicaliser à ce point les malades lorsqu'ils sont âgés : les interférences des médicaments n'améliorent pas leur état et sont même souvent nocives. Le troisième point concerne l'accès aux outils informatiques. Il faut faciliter l'accès à la domotique et intéresser les personnes âgées à Internet et à la télémédecine afin qu'ils continuent à avoir des activités. Enfin, il faut favoriser l'environnement domestique en plaçant, par exemple, les maisons de retraite au centre des villes ce qui favorise une socialisation accrue. En effet, il existe une corrélation entre l'environnement qui entoure la personne âgée (sanitaire, social et domestique) et son état de santé.

De plus, la prévention qui consiste à se maintenir en forme tout au long de sa vie est essentielle pour entamer la vieillesse dans de bonnes conditions. F. Forette, directrice de la Fondation Nationale de Gérontologie, précise aussi que la prévention de la dépendance se fait aussi par l'activité des seniors. L'activité est sûrement le meilleur moyen de conserver ses fonctions cognitives ainsi que son autonomie. Cela sans parler de l'avantage économique qu'il en résulte. Le statut d'inactivité d'un individu constitue aussi le meilleur indicateur de mortalité : de 35 à 80 ans, la mortalité est 3 fois supérieure chez les hommes inactifs, quelque soit la cause de l'inactivité.

Par conséquent, il existe de nombreux points porteurs d'espoir qui permettent de penser que la longévité peut encore augmenter. Sans en faire une liste non exhaustive, il est possible de réduire dès aujourd'hui les conduites à risque comme le tabagisme, l'alcoolisme, la violence routière, les accidents domestiques. De plus, il existe des systèmes en cours de diffusion comme les services de téléassistance et la domotique. Enfin, il reste des progrès attendus concernant les progrès médicaux et technologiques ainsi que l'éducation et la prévention des dépressions et des suicides.

5.2.8 Bilan sur les facteurs explicatifs de mortalité

Ces études menées par des économistes et/ou démographes permettent de sélectionner les facteurs explicatifs de mortalité. Nous considérons que les caractéristiques individuelles expliquant la mortalité sont l'âge, le sexe, la CSP, le statut matrimonial, le niveau d'éducation, le niveau de revenu ainsi que le niveau de dépendance qui reflète fortement le niveau de santé. Il est difficile d'avoir accès à ces données mais ce sont des facteurs à prendre en

compte pour estimer au mieux la mortalité individuelle.

De plus, il faut prendre en compte deux variables globales qui rendent compte de la région dans lequel évolue l'individu : le taux de chômage, un indicateur de l'offre de soin. Le taux de chômage (source INSEE) est un indicateur du niveau économique de la région. L'indicateur de l'offre de soins correspond par exemple à la densité de médecins dans la région ou à une information sur le secteur ambulatoire (CREDES, Base éco Santé). Si ces données ne sont pas disponibles, nous pouvons en tenir compte indirectement avec un indicateur des inégalités de revenus reflétant des inégalités socioéconomiques. Il existe de nombreux indicateurs d'inégalités des revenus comme l'indice de Gini, la mesure d'entropie de Theil, l'indice d'Atkinson ou tout simplement la variance du logarithme des revenus.

Enfin, il est possible de prendre en compte d'autres conditions environnementales et écologiques. Il faut faire attention à ne pas mettre trop de variables : Subramaniam et al. [34] ont montré que trop de variables explicatives provoquent une disparition de l'impact des inégalités.

5.3 Modélisation de la mortalité des individus par âge et par traits

Dans cette partie, nous décrivons un modèle de mortalité (de type logit) qui permet d'évaluer la probabilité de décès durant l'année en fonction de l'ensemble des caractéristiques (dont l'âge) de l'individu et de l'environnement dans lequel il évolue.

5.3.1 Description du modèle

Nous proposons une extension du modèle de Cairns classique en effectuant une régression logistique sur l'ensemble traits caractéristiques pour certaines tranches d'âge. L'intérêt de travailler par tranche d'âge est de ne pas avoir à considérer la corrélation entre les traits et l'âge en supposant que ces variables sont indépendantes sur une tranche d'âge.

Notons $\mathcal{A} = \{A_l, 1 \leq l \leq N\}$ la partition des classe d'âge avec N le nombre de tranches d'âge. Définissons $q(x, y, a, t)$ comme la probabilité de décès à la date t d'un individu de traits $x = (x_i)_{1 \leq i \leq M_x}$ d'âge a vivant dans un environnement (ou une région) de caractéristiques $y = (y_i)_{1 \leq i \leq M_y}$.

En fait, cette régression logistique sur chaque tranche d'âge peut s'écrire comme suit :

$\forall l \in \{1, \dots, N\}, \forall a \in A_l,$

$$\begin{aligned} \logit[q^1(x, y, a, t)] &= \log \left(\frac{q^1(x, y, a, t)}{1 - q^1(x, y, a, t)} \right) \\ &= \alpha_1(t) + a\alpha_2(t) + \sum_{i=1}^{M_x} \beta_i^l(t)^\top x_i + \sum_{j=1}^{M_y} \gamma_j^l(t)^\top y_j. \end{aligned}$$

Rappelons plus précisément la signification de chacune des variables et paramètres de ce modèle :

- Les variables y_j correspondent aux variables qui décrivent les caractéristiques globales comme l'offre de soins de la région ou les inégalités des revenus (indice de Theil ou variance du logarithme des revenus). On considère qu'il y a M_y variables globales à prendre en compte pour l'étude de la mortalité dans la région.
- Les variables x_i correspondent aux variables qui décrivent les caractéristiques individuelles explicatives (les traits) d'un individu parmi le sexe, le statut matrimonial, la CSP, le niveau de revenu, le niveau d'études, l'inactivité et la dépendance. Certaines de ces variables sont difficiles à déterminer lors du décès puisqu'elles ont évolué au cours de la vie de l'individu (CSP, revenu). On peut éventuellement ajouter certains traits qui affecteraient le déterminisme social et génétique [8] qui sont la CSP de la mère, la CSP du père, l'état de santé du père, et l'état de santé de la mère. On considère qu'il y a M_x variables individuelles à prendre en compte pour l'étude de la mortalité.
- On désigne par $N_{x,i}$ le nombre de classes au sein de la variable individuelle x_i et $N_{y,i}$ le nombre de classes au sein de la variable globale y_i . Il faut distinguer les variables quantitatives des variables catégorielles.

Dans le cas d'une variable quantitative x_i , cette variable ne possède qu'une seule classe $N_{x,i} = 1$ et prend ses valeurs dans un ensemble inclus dans \mathbb{R} .

Dans le cas d'une variable catégorielle, il existe plusieurs classes et la variable $x_i = (x_i^0, \dots, x_i^{N_{x,i}})$ prendra ses valeurs dans $\{0, 1\}^{N_{x,i}}$ où 0 signifie que l'individu n'appartient pas à cette classe et 1 qu'il y appartient.

Afin d'illustrer les notations, on propose deux exemples de variables.

La variable de sexe x_0 possède deux classes qui correspondent à "être une femme" et "être un homme" et $N_{x,0} = 2$. Le sexe est représenté par la variable catégorielle $x_0 = (1_{\{sexe=F\}}, 1_{\{sexe=H\}})$. Ainsi, la variable de sexe d'un homme sera $x_0^H = (0, 1)$ et la variable de sexe d'une femme sera $x_0^F = (1, 0)$.

Pour une variable x_1 correspondant au revenu, elle peut être décrite sous la forme d'une variable quantitative mais il est préférable de la décrire en quintiles avec $N_{x,1} = 5$. Les individus appartenant au premier quintile auront une variable $x_1^{1^{er} quintile} = (1, 0, 0, 0, 0)$, ceux du deuxième quintile auront la variable $x_2^{2^{e} quintile} = (0, 1, 0, 0, 0)$ et

ainsi de suite. D'ailleurs, il est possible de modéliser chacun des traits sous forme de variables catégorielles.

- Les tranches d'âge A_l sont à définir et on prendra des tranches de 5 ans par exemple avec les 25/30, 30/35 et ainsi de suite.

Remarque 5.6. *Pour la tranche des 0 à 25 ans, les traits n'impactent pas (ou très peu) le taux de mortalité par âge et on ne choisira peu de traits. De plus, les données pour cette tranche d'âge ne sont pas assez nombreuses du fait d'un nombre de décès très faible. Cela tombe bien puisqu'on considère que très peu de caractéristiques socioéconomiques ont de l'influence sur la mortalité aux âges faibles. Ainsi, on pourra modéliser la mortalité dépendant principalement de l'âge, du sexe et éventuellement des traits des parents afin de décrire les conditions de vies à la naissance.*

- $\alpha_1(t)$ et $\alpha_2(t)$ sont des coefficients réels reflétant l'influence d'une amélioration naturelle et d'une amélioration spécifique à l'âge.
- Les coefficients $(\beta_i^l)_{\substack{1 \leq i \leq M_x \\ 1 \leq l \leq N}}$ et $(\gamma_i^l)_{\substack{1 \leq i \leq M_y \\ 1 \leq l \leq N}}$ correspondent aux coefficients de la régression logistique, les premiers étant associés aux traits individuels et les autres aux variables globales d'environnement. L'indice l souligne le fait que chaque coefficient β_i^l ou γ_i^l est associé à une tranche d'âge A_l et sont constant sur cette tranche d'âge. En travaillant sur des petites tranches d'âge A_l , on suppose une indépendance de l'impact des traits par rapport à l'âge au sein d'une même tranche d'âge. Avec la définition des $N_{x,i}$, que la variable x_i soit catégorielle ou quantitative, le coefficient β_i^l est à valeurs dans $\mathbb{R}^{N_{x,i}}$ (un nombre réel pour une variable quantitative et un vecteur de taille $\mathbb{R}^{N_{x,i}}$ pour une variable catégorielle). Il en est de même pour les coefficients γ_i^l à valeurs dans $\mathbb{R}^{N_{y,i}}$. Ces coefficients vont permettre pour chaque tranche d'âge et chaque caractéristique de comparer l'impact d'une "classe" par rapport à une autre.

Le concept de la régression logistique présente l'avantage de prendre en compte les traits sous forme de variables catégorielles ou quantitatives. Cela donne la possibilité de comparer les catégories grâce à des outils comme celui de l'odd-ratio qui évalue le risque relatif d'une classe par rapport une autre. De plus, on peut effectuer des tests de significativité afin de savoir quelles variables sont explicatives de la mortalité et lesquelles ne le sont pas.

Avec un accès à un historique riche de données de mortalité par traits, il est possible de remplacer certains coefficients parmi les $\alpha_1(t)$, $\alpha_2(t)$, $(\beta_i^l)_{\{1 \leq i \leq M_x\}}$ et $(\gamma_i^l)_{\{1 \leq i \leq M_y\}}$ par des facteurs stochastiques de la même forme que dans le modèle de Cairns. On cherchera à estimer les paramètres grâce aux données historiques.

5.3.2 Méthode de calibration

Nous proposons une méthode de calibration permettant au modèle de rester consistant avec les courbes de mortalité par âge. En effet, les données disponibles pour l'ensemble des utilisateurs sont un historique de courbes de mortalité nationale par âge dans un pays pour chaque année. Il serait intéressant d'avoir un modèle qui ne change pas radicalement l'en-

semble des pratiques existantes en démographie, en assurance et dans d'autres domaines. Le fait d'étendre le modèle de Cairns à un ensemble de caractéristiques d'individus permettra de prendre en compte un nombre plus importants de facteurs de risques et permettra de réduire le risque de volatilité en considérant des informations sur l'individu autres que l'âge.

Ainsi, après avoir calibré le modèle standard de Cairns sur les tables de mortalité nationale, la problématique sera donc d'effectuer la régression logistique par tranche d'âge sur l'ensemble des traits explicatifs tout en conservant la calibration précédemment effectuée. Cela revient à faire une régression logistique conditionnellement à une autre régression logistique qui ne prend en compte que l'âge.

Analysons en théorie le procédé de calibration :

Tout d'abord, on calibre le modèle de mortalité par âge sur les tables de mortalité par âge classiques (mortality.org par exemple). Cela s'effectue de manière classique par la méthode des moindres carrés décrite en Section 5.1. Ainsi, on obtient les paramètres des facteurs $A_1(t)$ et $A_2(t)$ dont la dynamique est donnée par :

$$\begin{aligned} \text{logit} [q^0(t, a)] &= A_1(t) + aA_2(t), \\ A(t+1) &= A(t) + \mu + CZ_{t+1}. \end{aligned}$$

On considère le modèle général de mortalité par traits pour chaque tranche d'âge. Au sein d'une même tranche d'âge de 5 ans, l'influence de la caractéristique est supposée indépendante de l'âge. Dans le cas où une caractéristique x de l'individu n'est pas renseignée, la variable associée prend la valeur 0 si elle est quantitative et la valeur $\{0\}^{N(x)}$ si elle est catégorielle, de sorte que modèle coïncide avec le modèle de Cairns classique. En revanche, si la variable est renseignée, elle prend une valeur non nulle et modifie la valeur du taux de mortalité. De plus, on souhaite conserver la précision d'un modèle de mortalité par âge sans se restreindre à des tranches d'âge. Par conséquent, il faut que $\alpha_1(t) = A_1(t)$ et $\alpha_2(t) = A_2(t)$.

Ainsi, on a $\forall l \in \{1, \dots, N\}, \forall a \in A_l$

$$\begin{aligned} \text{logit}[q^1(y, x, a, t)] &= \log \left(\frac{q^1(y, x, a, t)}{1 - q^1(y, x, a, t)} \right) \\ &= A_1(t) + aA_2(t) + \sum_{i=1}^{M_x} \beta_i^l(t)^\top x_i + \sum_{j=1}^{M_y} \gamma_j^l(t)^\top y_j. \end{aligned}$$

Il est facile de voir que

$$\sum_{i=1}^{M_x} \beta_i^l(t)^\top x_i + \sum_{j=1}^{M_y} \gamma_j^l(t)^\top y_j = \text{logit} [q^1(y, x, a, t)] - \text{logit} [q^0(a, t)].$$

Ainsi, en considérant un autre ensemble de données de mortalité sur l'ensemble des caractéristiques pour chaque tranche d'âge, on obtient un historique dans le temps de données du type $\text{logit} [q^1(y, x, A_l, t)]$ pour chaque tranche d'âge et sur chaque caractéristique. Le terme $\sum_{i=1}^{M_x} \beta_i^l(t) x_i + \sum_{j=1}^{M_y} \gamma_j^l(t) y_j$ étant indépendant de l'âge au sein d'une tranche d'âge, on choisit a_l^* l'âge médian de la tranche d'âge A_l fournissant l'approximation suivante :

$$\sum_{i=1}^{M_x} \beta_i^l(t)^\top x_i + \sum_{j=1}^{M_y} \gamma_j^l(t)^\top y_j = \text{logit} [q^1(y, x, A_l, t)] - \text{logit} [q^0(a_l^*, t)] .$$

Le terme $\text{logit} [q^0(a_l^*, t)]$ correspond à l'historique de données de taux de mortalité (données nationales) par âge avec lequel on a calibré le modèle de Cairns (déjà calibré au préalable). Le terme $\text{logit} [q^1(y, x, A_l, t)]$ provient de la deuxième source de données disponible concernant la mortalité par tranche d'âge sur l'ensemble des caractéristiques étudiées. En considérant que les coefficients de la régression logistique pour chaque tranche d'âge β_i^l et γ_j^l qui sont eux même stochastiques. On décrit la dynamique des vecteurs $\beta^l = (\beta_1^l, \dots, \beta_{M_x}^l)^\top$ et $\gamma^l = (\gamma_1^l, \dots, \gamma_{M_y}^l)^\top$. On a $\forall 1 \leq l \leq N$, :

$$\begin{aligned} \beta^l(t) &= \beta^l(t-1) + \mu^{l,\beta} + C^{l,\beta} Z^\beta, \\ \gamma^l(t) &= \gamma^l(t-1) + \mu^{l,\gamma} + C^{l,\gamma} Z^\gamma, \end{aligned}$$

où $\mu^{l,\beta}$ (resp $\mu^{l,\gamma}$) est un vecteur de taille M_x (resp M_y) et $C^{l,\beta}$ (resp $C^{l,\gamma}$) est une matrice de corrélations entre les différentes caractéristiques de taille $M_x \times M_x$ (resp $M_y \times M_y$).

On pose $X(t) = \text{logit} [q^1(y, x, A_k, t)] - \text{logit} [q^0(a_k, t)]$ et on définit la série temporelle $\Delta X(t) = X(t) - X(t-1)$. Les coefficients $\mu^{l,\beta}$, $\mu^{l,\gamma}$, $C^{l,\beta}$ et $C^{l,\gamma}$ sont calibrés par la méthode de moindres carrés en utilisant les deux types de données disponibles. Dans la procédure de calibration, on associe les tables de mortalité nationale par âge utilisées habituellement aux informations issues de l'Échantillon Démographique Permanent de l'INSEE².

Par conséquent, ce modèle correspond à une extension du modèle de Cairns tout en conservant la calibration classique sur les tables de mortalité par âge. Il sera donc facile à utiliser par des compagnies d'assurance qui possèdent des données précises sur ses assurés. Enfin, ce modèle décrit l'évolution de l'impact des caractéristiques dans le temps en analysant par exemple l'évolution de l'impact d'une caractéristique sur la mortalité au cours du temps.

En résumé, ce modèle de mortalité étend la modélisation de Cairns décrivant la dynamique du logit du taux de mortalité par âge. Le facteur aléatoire est commun à tous les individus mais le taux de mortalité est individuel dans la mesure où il est fonction des traits caractéristiques de l'individu étudié. Malgré l'ajout de facteurs aléatoires, ce modèle

2. Nous remercions l'INSEE de nous avoir fourni ces données précieuses.

a l'avantage considérable de conserver le taux de mortalité positif puisque la fonction réciproque du "logit" prend des valeurs entre 0 et 1. En effet, on en déduit le taux de mortalité par âge et par traits $q(y_j, x_i, a)$ pour un ensemble de caractéristique individuelles et un ensemble de caractéristiques d'environnement pour chaque âge en inversant le logit par la relation suivante :

$$\begin{aligned} q(y_j, x_i, a) &= \frac{e^{\text{logit}[q(y_j, x_i, a)]}}{1 + e^{\text{logit}[q(y_j, x_i, a)]}} \\ &= \frac{1}{1 + e^{-\text{logit}[q(y_j, x_i, a)]}}. \end{aligned}$$

Enfin, le modèle étend un modèle standard de mortalité ce qui le rend plus transparent et facile à utiliser pour les acteurs de marché que ce soit en assurance ou en finance.

Remarque 5.7. *Ce modèle a en plus l'avantage de pouvoir corrélérer les facteurs de longévité à ceux que l'on pourrait utiliser afin de décrire l'évolution des taux d'intérêt. Ce modèle pourrait être utilisé afin d'évaluer des produits d'assurance vie qui mélangeraient risque de longévité et risque de taux d'intérêt.*

5.4 Étude quantitative de l'influence des caractéristiques sur la mortalité

Dans cette partie, nous considérons le modèle de mortalité par âge et par traits introduit dans la section précédente. Ce modèle tient compte d'un ensemble de caractéristiques des individus (dont l'âge) et va permettre d'analyser quel est l'impact de ces caractéristiques sur la mortalité ainsi que l'évolution de cet impact. Pour cela, il faut calibrer le modèle sur des données spécifiques.

5.4.1 Analyse de l'Échantillon Démographique Permanent

Nous avons été confronté au problème de confidentialité des données et il a été très difficile de nous les procurer. L'INSERM³ nous a fourni des données de mortalité par causes médicales de mortalité comme les cancers, les maladies cardiovasculaires, et autres causes. Ce sont des causes très importantes de mortalité mais ces données sont difficiles à exploiter et posent de nombreux problèmes. Tout d'abord, la cause "finale" du décès est renseignée par un médecin mais sa détermination n'est pas forcément fiable puisque de nombreux décès n'ont pas une cause unique et surviennent à la suite de la contraction de plusieurs maladies. De plus, il est difficile de savoir en quelle année un individu a réellement contracté la maladie puisqu'elle est souvent détectée à un stade avancé avec apparition de symptômes.

3. Nous remercions les chercheurs de l'INSERM et en particulier Grégoire Rey et Eric Jouglu avec qui nous avons eu des discussions intéressantes.

Il faudrait estimer chaque année la probabilité de décès d'un individu ayant une de ces maladies ce qui est loin d'être évident sans l'information concernant la date d'apparition de la maladie. Enfin, dans le but d'effectuer des scénarios de mortalité d'un individu, il faudrait aussi introduire des probabilités d'occurrence de ces maladies. Par conséquent, ce modèle est difficilement adaptable aux causes directes de mortalité comme les maladies car il faudrait des données précises sur l'occurrence de ces maladies.

Échantillon Démographique Permanent (EDP)

Ainsi, nous nous sommes intéressés aux données par caractéristiques en utilisant l'échantillon démographique permanent (EDP) de l'INSEE [16]. Ce fichier recense l'évolution de la situation de 992711 personnes (qui sont nées à partir de 1866) qui continuent à vivre actuellement ou qui sont décédées à une date (qui est renseignée). Ces personnes ont potentiellement été recensées à différentes dates 1968, 1975, 1982, 1990 et 1999 et nous avons accès à différentes caractéristiques pour des personnes qui sont toujours vivantes et qui ont rempli leur fiche de recensement. Cet échantillon contient l'ensemble des français nés un jour parmi quatre jours de l'année (1, 2, 3 et 4 octobre) et contient un dossier individuel avec l'ensemble des bulletins de recensement de 1968 à 1990 ainsi que les bulletins statistiques d'état civil et les mentions de décès. Le fait de suivre la population va permettre d'étudier un taux de mortalité non pas en fonction d'une situation figée (avec la déclaration au moment du décès) mais en fonction de la situation à différentes dates. Ainsi, il sera possible de tenir compte d'éventuels changements de situation.

En effet, un problème classique est de mesurer l'impact d'une caractéristique qui évolue dans le temps. Si nous cherchons à mesurer la mortalité par milieu social, une méthode classique qui a longtemps été utilisée était d'analyser la catégorie sociale au moment du décès. Pour la mesurer, il suffisait de considérer pour chaque catégorie sociale le ratio des décès de l'année d'un recensement (ou sur une période contenant le recensement) divisé par l'effectif de la population. Cette méthode est assez simple mais conduit à des résultats parfois surprenants résultant de la distorsion entre numérateurs et dénominateurs. En effet, les informations collectées au décès et les informations issues de recensement sont faites dans des conditions différentes et sont traitées différemment. Dès 1907, M. Huber analysait ce problème : la situation professionnelle étant mouvante et pouvant concerner des métiers mal identifiés, le risque de distorsion devient important. Et il est d'ailleurs possible que la déclaration du bulletin de décès ait tendance à valoriser la profession du défunt en rehaussant son statut.

Pour résoudre ce type de problème, l'INSEE a mis en place une observation longitudinale consistant à suivre des individus dont les caractéristiques sont recensées à différentes dates

pour lesquels ils renseignent la date de décès. Cela permet de passer outre les problèmes de distorsion en ayant un suivi de la catégorie sociale de l'individu au cours de sa vie. Concernant les individus de l'Échantillon Démographique Permanent qui sont partis vivre à l'étranger, le consulat en informe l'INSEE qui met à jour son répertoire. Ainsi, les distorsions entre les déclarations de l'état civil et les recensements sont supprimées. Le cas des retraités peut facilement être traité car nous avons un suivi des catégories socioprofessionnelles qu'ils ont connu durant leur vie et nous choisissons la catégorie "représentative". D'autre part, cette source de donnée ne vieillit pas dès lors que l'échantillon est renouvelé (avec les nouvelles naissances) et mis à jour à chaque recensement.

Concernant la significativité de certaines caractéristiques, de nombreuses études ont été menées sur la population française et l'échantillon démographique permanent expliquant que la CSP, le statut matrimonial, le niveau d'études ont une influence sur la mortalité (voir Section précédente). Nous ne reproduirons pas ces tests qui consistent à effectuer des régressions logistiques et à faire des tests de significativité en utilisant les odd ratio. Ces études ont été menées par de nombreux chercheurs dont F. Jusot au cours de son doctorat [17].

Étude de la précision des données de l'EDP

On s'intéresse maintenant aux données de l'EDP. Tout d'abord, on vérifie que cet échantillon est représentatif de la population au sens où les taux de mortalité par âge et par date sont à peu près les mêmes. On calcule le taux de mortalité $q^{EDP}(A_l, t)$ à une date t sur une tranche d'âge A_l obtenu avec l'échantillon démographique et on le compare au taux de mortalité national $q(a_l^*, t)$ (provenant de la Human Mortality Database (HMD)) à l'âge médian a_l^* . Par exemple, on compare la probabilité d'un individu de l'EDP dont l'âge appartient à la tranche d'âge 40/45 de décéder au cours de l'année 1980 au taux de mortalité nationale à 42 ans observé cette même année. Les probabilités se révèlent être très proches les unes des autres mais elles commencent à s'éloigner les âges élevés. Cela résulte du manque de données dans l'échantillon mais aussi des différences de niveau de mortalité au sein d'une même tranche d'âge. On peut donner une borne supérieure heuristique qui est toujours vérifiée empiriquement sur ces données.

$$q^{EDP}(A_l, t) - q(a_l^*, t) < 0.1 \times q^{EDP}(A_l, t).$$

On a d'ailleurs comparé les logit de mortalité (changement d'échelle) pour avoir un point de vue plus précis sur les différentes tranches d'âge. On obtient pour les hommes les résultats suivants (même type de résultats pour les femmes) :

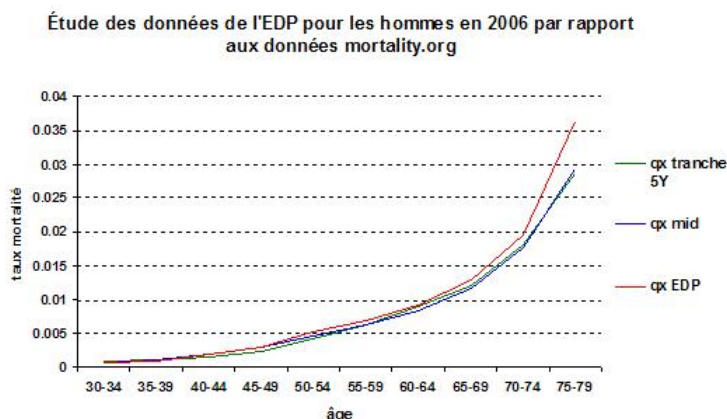


FIGURE 5.30 – Précision des données de mortalité pour les hommes issues de l'EDP par rapport aux tables de mortalité par tranche d'âge ou pour l'âge médian

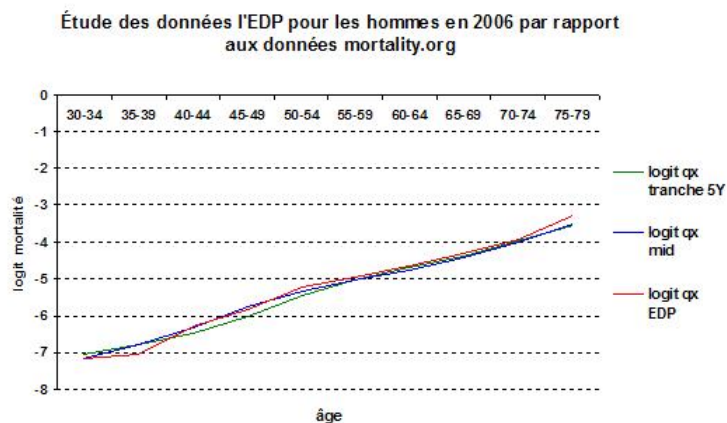


FIGURE 5.31 – Précision des données en logit de mortalité pour les hommes issues de l'EDP par rapport aux tables de mortalité par tranche d'âge ou pour l'âge médian

Ces résultats sont valables pour l'ensemble des tranches d'âge (tranches de 5 ans pour les personnes âgées de 25 à 85 ans) que l'on a étudié et sur l'ensemble des dates de 1968 à 1990. On peut en conclure que l'échantillon démographique est représentatif de la population française et cela confirme aussi que l'approximation de l'âge médian pour une tranche d'âge est très bonne. Pour information, l'écart entre le logit de mortalité des données de l'EDP par âge et le logit de mortalité des données nationales reste relativement faible sur l'ensemble des tranches d'âge et sur chaque date. Cela prouve bien que la fonction logit a de bonnes propriétés intrinsèques pour l'étude des taux de mortalité. En effet, la fonction logit permet d'analyser les taux de mortalité à différentes tranches d'âges avec une erreur de

précision du même ordre de grandeur quelque soit l'âge. Il serait intéressant de comprendre si il existe une cause liée à la structure de l'échantillon proposé autre que sa taille qui est plus faible que la taille de la population nationale.

Maintenant qu'on sait que les calculs de probabilités de décès sont raisonnablement précis pour chaque tranche d'âge, on va chercher à calculer des taux de mortalité par tranche d'âge et par traits. On considère l'échantillon de la population à une date donnée et pour chaque tranche d'âge et on distingue les individus ayant une caractéristique donnée. Cela permet de calculer un taux de mortalité pour chaque tranche d'âge et chaque trait.

Commentaires sur la calibration du modèle sur les données

L'objectif est d'effectuer des projections à l'aide du modèle de mortalité par âge et par traits. Pour cela, on analyse l'influence des différentes caractéristiques sur un historique de données, on calibre les paramètres du modèle et on en déduit des scénarios d'évolution de la mortalité en fonction de différentes caractéristiques. On rappelle que la modélisation est une régression logistique indirecte puisqu'en pratique on aboutit à un modèle logistique mais avec une procédure légèrement différente. En effet, on n'effectue pas la régression logistique directement, car on a souhaité être cohérent avec les modèles de mortalité par âge sans perdre de précision due à la modélisation par tranche d'âge. De plus, si on considérait directement la régression logistique classique par âge et par traits et qu'on ne renseignait pas les traits des individus, on ne pourrait pas récupérer les courbes nationales de mortalité par âge. Cela s'explique par le fait que les paramètres d'une régression logistique dépendent du flux d'information en entrée et qu'il est nécessaire de renseigner tous les traits de l'individu si on veut utiliser le modèle.

On souligne une propriété importante sur la précision des données de l'EDP. Le premier recensement commence en 1968 mais les données de l'EDP deviennent de plus en plus précises dans le temps avec un plus grand nombre de personnes recensées. Ainsi, il faut trouver un bon compromis sur le nombre de données à prendre : il faut prendre en compte des données précises mais il faut aussi estimer les paramètres du modèle pour lequel un nombre suffisant de données est nécessaire. Il n'est pas possible de faire un "backtesting" car il n'existe pas de tables de référence pour les taux de mortalité par âge et par traits. Le modèle de mortalité par âge utilisé sera en revanche calibré en ayant fait une procédure de backtesting empirique (voir Section précédente) au préalable afin de déterminer les sets de calibrations adaptés.

Enfin, il est évidemment possible d'améliorer le traitement des données par de nombreuses méthodes statistiques afin d'obtenir des calculs de probabilités plus précis. Nous évitons au maximum les données manquantes (en ne les prenant pas en compte), mais il est sûrement possible d'améliorer ces résultats avec des méthodes statistiques plus robustes. En effet,

le fait de ne pas considérer ces données manquantes restreint le nombre de personnes de l'échantillon et rend les calculs de probabilités moins précis. De plus, il est assez difficile d'obtenir des résultats précis sur la mortalité et de nombreuses études sont faites pour améliorer le calcul des probabilités de décès.

Le but de cette étude n'étant pas de fournir des résultats parfaits ou de référence mais plutôt de comprendre comment calibrer le modèle de mortalité par âge et par traits et de comprendre quels type de développements et résultats peuvent être mis en place. Les résultats trouvés sont souvent assez intuitifs, nous laissons aux personnes intéressées le soin d'implémenter leur propre modèle et leur propre méthode de calibration.

Informations complémentaires sur les données de l'EDP

Pour avoir une idée de la taille de l'échantillon utilisé dans le calcul des probabilités de décès, on donne un exemple du type d'information que l'on peut récupérer avec ces données. On étudie l'influence du statut matrimonial sur la mortalité en 2003. On peut recenser le nombre de personnes vivantes et le nombre de décès pour chaque tranche d'âge en distinguant le sexe et le statut matrimonial de l'individu.

âge	non recensés décès	non recensés vivants	mariés décès	mariés vivants	non mariés décès	non mariés vivants
25/30	0	2967	2	1085	14	22146
30/35	1	2684	2	6717	23	19438
35/40	2	1839	14	12206	33	15093
40/45	1	1437	22	15144	32	11422
45/50	1	1034	46	16631	58	8790
50/55	1	1077	74	18132	79	7180
55/60	1	977	91	16437	68	5095
60/65	2	659	104	12095	69	3495
65/70	2	480	171	12610	44	2892
70/75	0	269	224	11326	90	2623
75/80	0	152	328	8508	142	2132
80/85	0	101	366	5693	132	1679
85et+	0	136	373	7398	249	4160

TABLE 5.1 – Tableau de la structure des hommes l'échantillon en 2003 permettant le calcul des probabilités de décès par statut maritaux

âge	non recensées décès	non recensées vivantes	mariées décès	mariées vivantes	non mariées décès	non mariées vivantes
25/30	1	2964	1	2861	7	19721
30/35	0	2558	2	9678	15	16233
35/40	1	1781	1	14298	8	12478
40/45	0	1194	15	16541	13	10013
45/50	0	1095	22	17178	22	8276
50/55	0	969	36	17784	30	7472
55/60	2	741	34	15194	32	5737
60/65	1	492	44	12002	28	4277
65/70	0	346	66	11628	54	4919
70/75	0	267	156	10201	91	6277
75/80	1	173	156	7601	192	7142
80/85	0	162	167	4436	299	6737
85 et +	2	215	194	4202	1227	17175

TABLE 5.2 – Tableau de la structure des femmes l'échantillon en 2003 permettant le calcul des probabilités de décès par statut maritaux

Tout d'abord, concernant le statut matrimonial, on observe que l'effectif des femmes aux âges très élevés est près de deux fois plus élevé que celui des hommes et qu'il y a beaucoup plus de femmes veuves que d'hommes veufs. Cette information est souvent soulignée par les démographes lorsqu'ils publient des analyses de la population française. Cela traduit bien le fait que les femmes vivent plus longtemps que les hommes et que leur mari décède souvent en premier.

Remarque 5.8. *On peut voir tout d'abord que l'échantillon est assez important mais que ça n'est pas encore idéal pour un calcul précis de probabilités de décès. Il faudrait avoir un échantillon plus important. L'autre alternative est de ne pas considérer uniquement les décès d'une année mais plutôt l'ensemble des décès sur une période donnée. Par exemple, au lieu d'avoir 5 taux de mortalité entre 2000 et 2004, on aurait un taux de mortalité constant pour la période 2000/2004. Il faudrait adapter le modèle d'évolution à ce cadre pour avoir une évolution du taux entre chaque période.*

5.4.2 Étude de l'influence de la catégorie socioprofessionnelle

On cherche à analyser le taux de mortalité par catégories socioprofessionnelles (CSP), à mettre en place le modèle de mortalité et à le calibrer. Chaque année, le nombre d'individus décédés par âge et par CSP est rapporté au nombre d'individus vivants durant cette année avec ses mêmes caractéristiques. Tout d'abord, on propose une première classification des catégories socioprofessionnelles :

- inactif : 0
- agriculteurs : 1
- artisans, commerçants, chef d'entreprises : 2
- cadres et professions intellectuelles supérieures : 3
- Professions intermédiaires : 4
- Employés : 5
- Ouvriers : 6

On a choisi de prendre cette classification en regardant des études faites par l'INSEE qui publie un indice standardisé de mortalité (ISM) par CSP. Cet indice est un indice multiplicatif de mortalité par rapport à une mortalité de référence (celle des employés) ce qui permet de comparer sur une période la mortalité entre différentes catégories socioprofessionnelles (voir Table 5.3). L'INSEE a calculé ces indices de mortalité en analysant la mortalité des personnes d'âge compris entre 45 ans et 64 ans sur les dates de 1982 à 2001.

	Femmes	Hommes
N'a jamais travaillé	1.2	1.9
Travaille ou a déjà travaillé	1	1
Cadres ou professions intellectuelles supérieures	0.8	0.6
Professions intermédiaires	0.8	0.9
Artisans, commerçants, chef d'entreprise	0.9	0.9
Employés	1.0	1.0
Agriculteurs exploitant	0.9	0.8
Ouvriers	1.1	1.2

TABLE 5.3 – Indice standardisé de mortalité par CSP et par sexe sur la période de 1982 à 2001 pour des hommes et femmes de 45 à 64 ans

L'analyse de ce tableau permet de souligner plusieurs phénomènes. Tout d'abord, la classe des inactifs connaît une surmortalité très importante, surtout chez les hommes avec une surmortalité de 90% par rapport à la mortalité moyenne des hommes. De plus, il existe des classes protectrices comme celle des cadres et professions intellectuelles supérieures alors que certaines classes sont exposées à une surmortalité comme celle des ouvriers. Enfin, la dernière chose importante à relever est que la structure socioprofessionnelle a davantage d'effets sur la mortalité des hommes que sur la mortalité des femmes et que l'inactivité a un impact moindre sur les femmes. Cela nous amène donc à penser que nos résultats seront sûrement significatifs pour les hommes alors que pour les femmes, ils seront plus difficiles à interpréter.

On ne peut pas considérer les données pour chaque âge, chaque CSP et chaque année puisque les probabilités calculées seraient trop imprécises. En modélisant l'ensemble des âges entre 20 et 90 ans et l'ensemble des CSP qui sont au nombre de 7 dans la classification globale, il faudrait donc recenser $70 \times 7 = 560$ types d'individus et ce qui revient à calculer 560 probabilités de décès à chaque date. Le calcul se révélerait imprécis et fournirait des résultats invraisemblables. Il faut donc trouver un moyen pour regrouper certaines catégories d'individus afin d'obtenir des probabilités de décès assez précises. On note que la méthode nécessite de plus une indépendance de l'impact des traits et de l'âge car le modèle postule que l'évolution de la mortalité par âge est indépendante de celles des caractéristiques.

Un moyen de passer outre ce problème, tout en conservant une indépendance entre l'évolution par âge et par traits, est de considérer un modèle de mortalité par tranche d'âge. Il faut étudier la mortalité par caractéristiques et par tranche d'âge, ces tranches d'âge ne devant pas être trop "larges". Pour l'ensemble des personnes d'âge compris entre 25 ans et 90 ans, on obtient 14 tranches d'âge de 5 ans.

De plus, le calcul à chaque année de l'ensemble des probabilités par tranche d'âge et par CSP correspond à $13 \times 7 = 91$ probabilités à calculer pour chaque date. Si on estime ces probabilités, on se rend compte que les résultats restent encore imprécis. Pour une étude pratique, il faut encore affiner ce calcul de probabilité. En effet, l'idée de la modélisation va être de regarder l'évolution de l'impact des CSP dans le temps (sur une durée de 38 ans environ de 1968 à 2006) et de calibrer les paramètres pour qu'il soit cohérent avec l'historique des variations. Si de la date t à $t + 1$, il y a une imprécision trop importante due à un calcul pas assez précis des probabilités de décès, le modèle n'aura pas grand intérêt et ne fournira pas des résultats satisfaisants. Pour avoir un calcul plus précis de probabilités de décès par caractéristiques, on a analysé l'ISM (voir Table 5.3) qui incite à rassembler les CSP dans des classes plus larges de la manière suivante :

- inactifs (hors retraités) et non déclarés (données non exploitables) : CSP_0
- catégories "supérieures" (cadres et professions intellectuelles supérieures, artisans, commerçants, chefs d'entreprise, professions intermédiaires, agriculteurs) : CSP_1
- catégories "inférieures" (employés et ouvriers) : CSP_2

Cela amène à définir la variable CSP comme une variable catégorielle possédant 3 classes :

$$CSP = (1_{\{CSP=CSP_0\}}, 1_{\{CSP=CSP_1\}}, 1_{\{CSP=CSP_2\}}).$$

Par conséquent, on peut écrire le modèle de mortalité par âge et trait de la manière suivante :

$$\begin{aligned} \forall a \in A_l, \quad \text{logit}(q(CSP, a, t)) = & A_1(t) + aA_2(t) + \beta_{CSP,0}^l(t)1_{\{CSP=CSP_0\}} \\ & + \beta_1^l(t)1_{\{CSP=CSP_1\}} + \beta_2^l(t)1_{\{CSP=CSP_2\}}. \end{aligned}$$

La somme des indicatrices est égale à 1 et correspond à la variable intercept sous jacente à A_1 . Les études de l'INSEE ainsi que l'ISM révèlent que la catégorie "supérieure" doit

avoir une mortalité inférieure à celle de la catégorie "inférieure" et une mortalité encore inférieure à la catégorie des inactifs. Les inactifs ont un statut à part car ils connaissent une surmortalité très importante par rapport à n'importe quelle autre classe professionnelle. Le problème est qu'il est difficile de distinguer dans les données de l'EDP les personnes inactives des personnes n'ayant pas déclaré leur situation professionnelle au cours des différents recensements. On a choisi de rassembler ces deux types de personnes au sein d'une même classe qui sera difficile à analyser. Il y aura bien évidemment un biais lié aux déclarations erronées mais il est naturel d'obtenir la classification suivante :

$$\forall t, \forall A_l, \quad q(CSP_0, A_l, t) \geq q(CSP_2, A_l, t) \geq q(A_l, t) \geq q(CSP_1, A_l, t).$$

On vérifie cela assez facilement sur l'ensemble des données puisque sur chaque tranche d'âge, la probabilité nationale de décès est presque toujours comprise entre celle de la catégorie supérieure et celle de la catégorie inférieure. Il est possible d'analyser les différences liées aux caractéristiques sur les logit de mortalité ou sur le taux de mortalité directement. En effet, la modélisation repose sur des différences de logit de mortalité et il faut analyser ces différences que l'on injecte dans notre modèle. Cependant, la fonction logit n'étant pas une fonction linéaire, il peut être difficile d'interpréter les écarts de mortalité. On représente les écarts en taux de mortalité mais qui sont difficiles à interpréter du fait d'une grande variation suivant les âges. On prend l'exemple de la tranche d'âge des 65/69 ans et on représente l'évolution dans le temps de la surmortalité de la classe inférieure par rapport à la moyenne nationale.

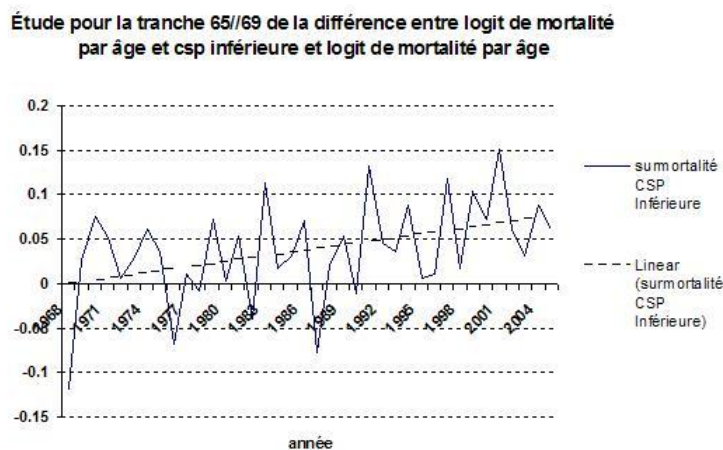


FIGURE 5.32 – Évolution du facteur $\beta^{CSP_2}(t)$ entre 1968 et 2005

Les données de l'EDP révèlent une surmortalité de la classe inférieure par rapport à la mortalité moyenne ainsi qu'une sous mortalité de la classe supérieure par rapport à la mortalité moyenne. Il est possible d'essayer de capter une tendance d'évolution de cette surmortalité

et de cette sous mortalité en considérant une période pendant laquelle on estime qu'il n'y a pas eu de rupture de tendance. Cependant il apparaît difficile de capter une volatilité car les données sont très bruitées. En effet, les variations observées ne correspondent pas un effet de volatilité mais plutôt à un bruit important des données du fait d'un nombre peu précis de données. En pratique, il faudrait un échantillon plus important afin de pouvoir estimer une véritable volatilité d'évolution ce qui limite à l'étude d'une tendance d'évolution. Concernant la différence de mortalité initiale, on fait une moyenne sur les 5 dernières années entre 2000 et 2005 et on obtient les courbes de mortalité suivantes :

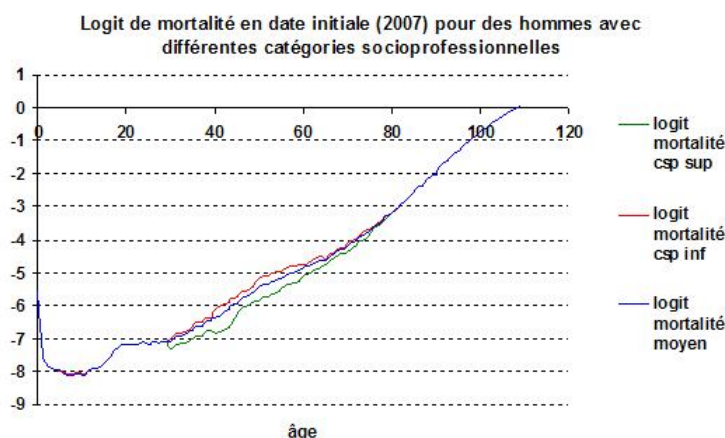


FIGURE 5.33 – Logit de mortalité des hommes pour différentes CSP calculés avec les données de 2000 à 2005

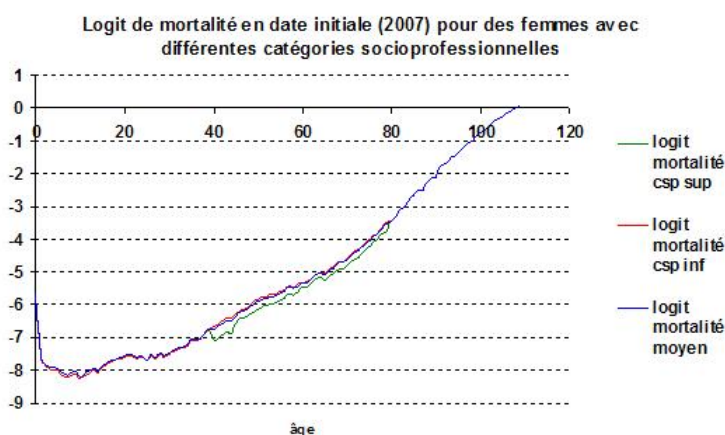


FIGURE 5.34 – Logit de mortalité des femmes pour différentes CSP calculés avec les données de 2000 à 2005

La mortalité des hommes est tout à fait cohérente avec ce que l'on attend mais que la catégorie sociale des femmes est difficile à prendre en compte. Cela s'explique par le fait que la structure de catégories socioprofessionnelles affecte moins les femmes que les hommes. Ayant calibré au préalable le modèle de Cairns par âge et utilisant les informations obtenues sur la différence de mortalité par CSP (différence initiale et tendance), on a calibré un modèle d'évolution de mortalité par âge et CSP. En calibrant les effets de tendances, on peut essayer d'analyser la déformation des tables de mortalité dans le temps. Ainsi, on représente la mortalité moyenne prévue pour 2017 :

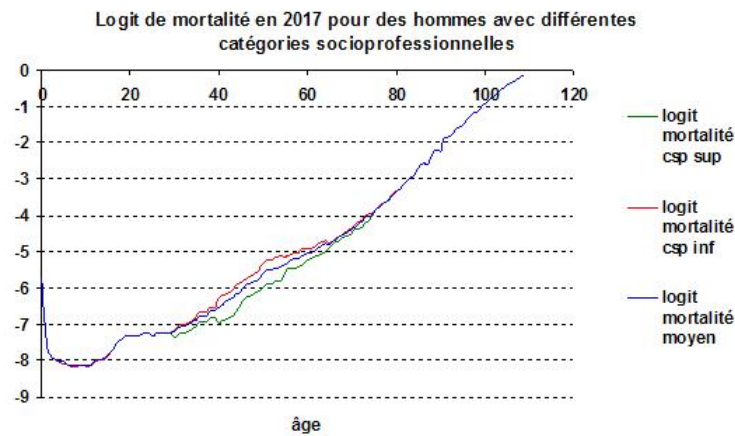


FIGURE 5.35 – Courbes de logit de mortalité des hommes en 2017 pour différentes CSP et par âge

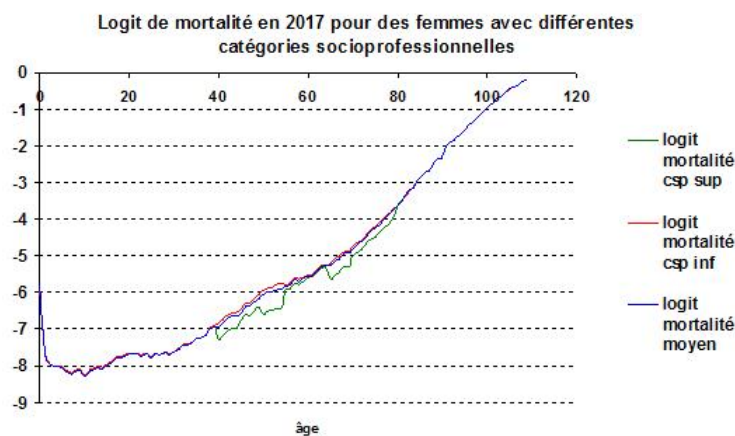


FIGURE 5.36 – Courbes de logit de mortalité des femmes en 2017 pour différentes CSP et par âge

Ces courbes soulignent plusieurs phénomènes. Tout d'abord, la CSP a un réel impact sur la mortalité des hommes alors que les femmes sont bien moins affectées par la catégorie sociale. Si on analyse les écarts de mortalité avec la mortalité moyenne, on remarque que l'impact de la catégorie supérieure a tendance à diminuer alors que l'impact de la catégorie inférieure a tendance à augmenter. Cela peut s'expliquer par le fait que les hommes appartiennent de plus en plus à la catégorie sociale supérieure et de moins en moins dans la catégorie inférieure ce qui rend la catégorie inférieure de plus en plus discriminante par rapport à la catégorie moyenne.

Ainsi, le problème de l'étude de l'impact de la CSP est assez compliqué pour plusieurs raisons. Tout d'abord, les femmes apparaissent bien moins affectées par la structure des catégories socioprofessionnelles et cette variable se révèle par conséquent difficile à étudier et à interpréter concernant le sexe féminin. De plus, la CSP est une variable changeante dans le temps et bien que l'on ait fait une étude longitudinale adaptée à ce genre de variable, il n'est pas facile de capter l'évolution de l'impact de cette variable. Il faudrait des variables moins changeantes dans le temps comme le niveau de revenu (dont l'évolution est stable dans le temps) ou le niveau de diplôme.

Remarque 5.9. *La classe des catégories socioprofessionnelles reflète généralement la situation et les conditions de vie d'un individu. On aurait pu analyser l'impact du niveau de revenu qui est une variable purement économique et qui est significative. Le niveau de revenu dont l'évolution est assez stable dans le temps est connu pour avoir un impact majeur dans le niveau de mortalité [17] et donnerait des résultats sûrement intéressants. Les données sur les niveaux de revenu n'étant pas accessibles, on analyse quelle est l'influence du niveau de diplôme qui est une autre variable socioéconomique.*

5.4.3 Étude de l'influence du niveau de diplôme

Cette partie concerne l'analyse de la variable du niveau d'éducation au travers du niveau de diplôme. Cette variable, n'ayant pas évolué significativement dans le temps puisque les diplômes sont les mêmes depuis près de 50 ans, l'évolution sera plus facile à analyser. On a choisi de prendre comme niveau de référence le diplôme du baccalauréat qui est la référence nationale. On distingue d'une part les personnes ayant poursuivi leurs études avec l'obtention d'un diplôme universitaire et d'autre part les personnes qui n'ayant pas été au delà du baccalauréat.

- sans diplôme ou non déclarés (données non exploitable) : DIP_0
- diplôme supérieur ($> \text{BAC}$) : DIP_1
- diplôme inférieur ($\leq \text{BAC}$) : DIP_2

La variable de diplôme est définie par

$$DIP = (1_{\{DIP=DIP_0\}}, 1_{\{DIP=DIP_1\}}, 1_{\{DIP=DIP_2\}}).$$

$$\forall a \in A_l, \quad \text{logit}(q(DIP, a, t)) = A_1(t) + aA_2(t) + \beta_{DIP,0}^l(t)1_{\{DIP=DIP_0\}} + \beta_{DIP,1}^l(t)1_{\{DIP=DIP_1\}} + \beta_{DIP,2}^l(t)1_{\{DIP=DIP_2\}}.$$

Comme pour la CSP, on peut analyser les facteurs β_i^l et voir leur évolution dans le temps. Quelque soit le sexe considéré, il existe un différentiel de mortalité assez important entre les personnes ayant un diplôme supérieur et les personnes n'en ayant pas. Cela se vérifie chez les hommes et chez les femmes dont la mortalité est significativement affectée par cette nouvelle variable.

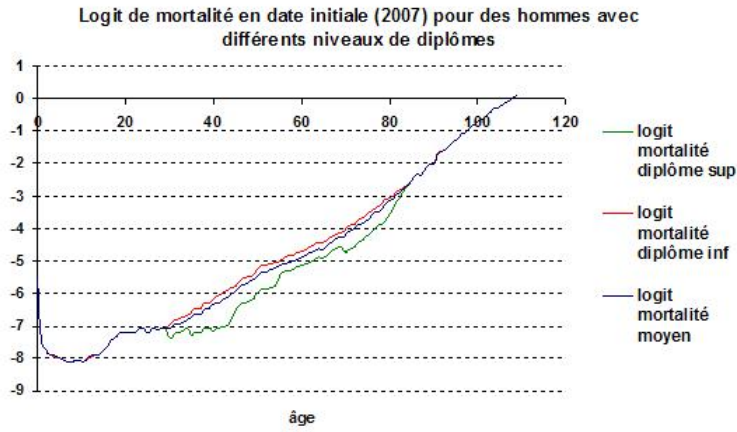


FIGURE 5.37 – Logit de mortalité des hommes pour différents niveaux de diplôme calculés avec les données de 2000 à 2005

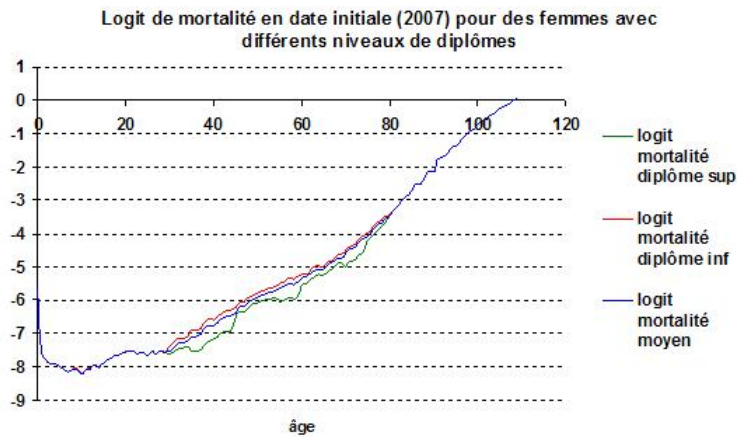


FIGURE 5.38 – Logit de mortalité des femmes pour différents niveaux de diplôme calculés avec les données de 2000 à 2005

Ainsi, on peut effectuer une projection tendancielle afin de produire des tables de mortalité prospectives par âge et par niveau de diplôme. On représente l'évolution moyenne de la mortalité en 2017 pour les hommes et les femmes.

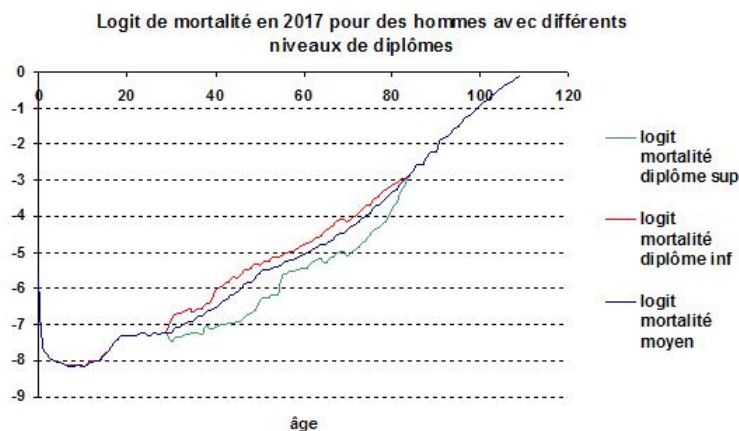


FIGURE 5.39 – Courbes de logit de mortalité des hommes en 2017 pour différents niveaux de diplômes et par âge

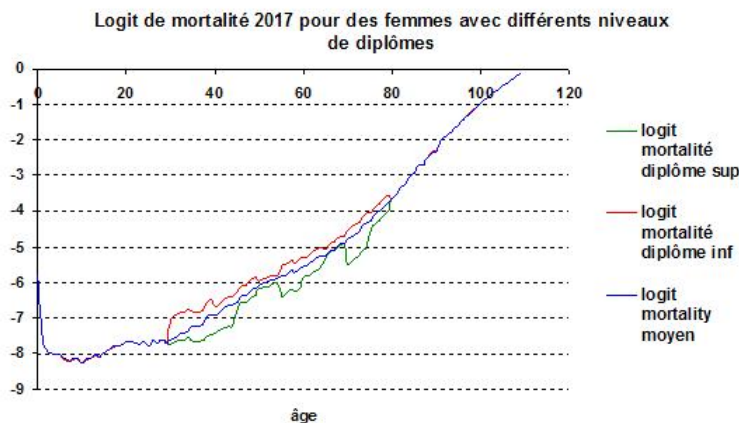


FIGURE 5.40 – Courbes de logit de mortalité des femmes en 2017 pour différents niveaux de diplômes et par âge

Ces graphiques confirment le fait que l'influence du niveau de diplôme est très importante et qu'elle affecte à la fois l'homme et la femme sur l'ensemble des tranches d'âge. Le fait d'avoir un diplôme universitaire procure une baisse assez importante du taux de mortalité par rapport au niveau de mortalité moyen. Par conséquent, les études universitaires "protègent" contre la mortalité. On remarque aussi que l'écart de mortalité entre les personnes

diplômées et les personnes non diplômées va avoir tendance à augmenter dans le temps sur l'ensemble des tranches d'âge. Enfin, il est inutile de pousser notre étude aux âges très élevés (supérieur à 85 ans) car l'influence du niveau de diplôme s'estompe et que les données sont peu nombreuses. On souligne aussi un impact tout de même significatif pour les personnes âgées entre 70 et 80 ans.

La probabilité de décès par diplôme a connu des variations assez complexes ces derniers temps. Depuis une quinzaine d'années, de plus en plus d'étudiants obtiennent le diplôme du baccalauréat ce qui va affecter l'importance de ce diplôme. Il est difficile de prédire son évolution dans les prochaines années. La variable du diplôme est tout de même plus stable dans le temps que les catégories socioprofessionnelles et le niveau affecte significativement l'homme et la femme. Ainsi, il est sûrement préférable de considérer cette variable socioéconomique qui est explicative de la mortalité à la fois pour les hommes et les femmes.

5.4.4 Étude de l'influence du statut matrimonial

L'influence de la vie de famille est assez complexe pour une autre raison. En effet, le changement actuel des mœurs pose des problèmes de modélisation de cette variable. On a longtemps considéré le mariage protecteur mais avec la conjoncture actuelle, les démographes s'accordent à dire qu'il apparaît plus concret de dire que c'est la vie familiale qui est protectrice. Actuellement, de nombreuses personnes sont en couple en étant en Pacs ou en ménage libre sans être mariées officiellement. D'ailleurs, la moitié des bébés naissent hors mariage [23]. D'un autre côté, des études révèlent que la vie en couple sans mariage protège moins que le mariage [31]. Par conséquent, le choix de la bonne variable qui transcrit la notion de vie de famille est difficile à déterminer.

On pourrait étudier l'impact de la vie familiale qui est plus conservateur que le mariage. Ici, en l'absence de cette information, on s'intéresse uniquement au mariage. On compare le statut "marié" et "non marié" mais il faudra garder à l'esprit que de nombreux célibataires peuvent avoir une vie familiale.

On distingue 3 classes concernant le statut matrimonial

- non déclarés (données non exploitable) : SM^0
- mariés : SM^1
- célibataires, divorcés (sans remariage) ou veufs (sans remariage) : SM^2

Dans ce cadre, la variable du statut matrimonial se définit par

$$SM = (1_{\{SM=SM_0\}}, 1_{\{SM=SM_1\}}, 1_{\{SM=SM_2\}}) .$$

Le modèle peut se réécrire classiquement sous la forme

$$\forall a \in A_l, \quad \text{logit}(q(SM, a, t)) = A_1(t) + aA_2(t) + \beta_{SM_0}^l(t)1_{\{SM=SM_0\}} + \beta_{SM_1}^l(t)1_{\{SM=SM_1\}} + \beta_{SM_2}^l(t)1_{\{SM=SM_2\}}.$$

On peut encore voir les écarts de mortalité entre les différents statuts pour l'ensemble des tranches d'âge, et voir son évolution tendancielle dans les années à venir. Actuellement, il existe une différence significative de mortalité pour une personne mariée et pour une personne non mariée quelque soit son âge. Cela peut se voir sur les graphiques suivants :

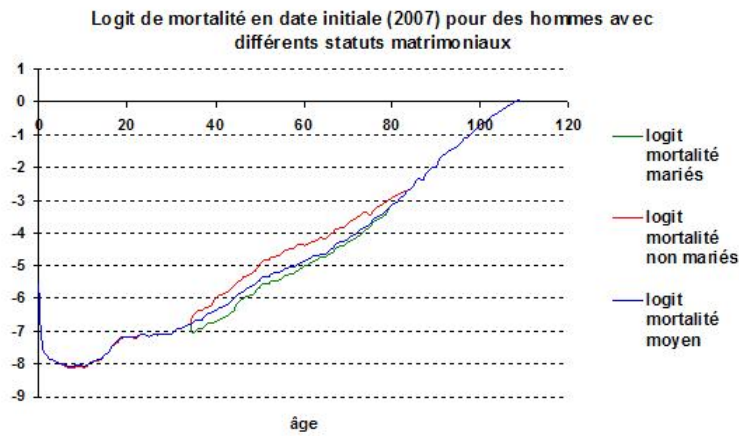


FIGURE 5.41 – Logit de mortalité des hommes pour différents statuts maritaux calculés avec les données de 2000 à 2005

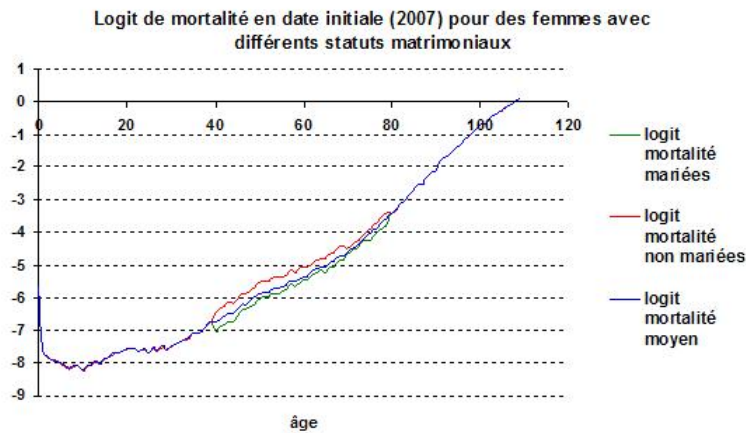


FIGURE 5.42 – Logit de mortalité des femmes pour différentes statuts maritaux calculés avec les données de 2000 à 2005

On peut aussi analyser la déformation de ces courbes de logit de mortalité dans le temps ainsi que leurs projections moyennes en 2017 calculées grâce au modèle de mortalité par âge et par traits.

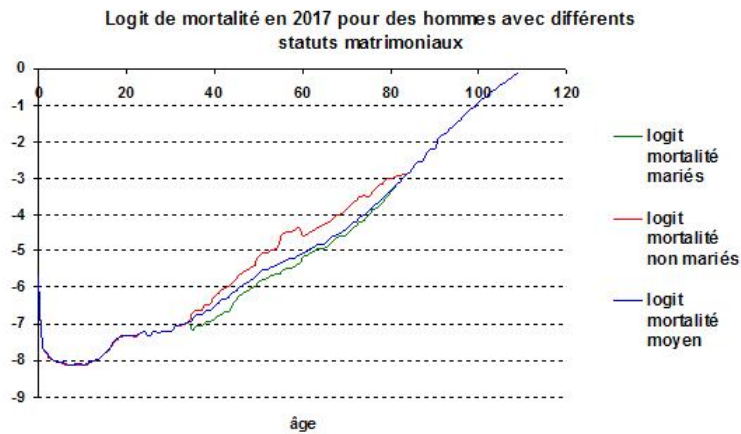


FIGURE 5.43 – Courbes de logit de mortalité des hommes en 2017 pour différents statuts maritaux et par âge

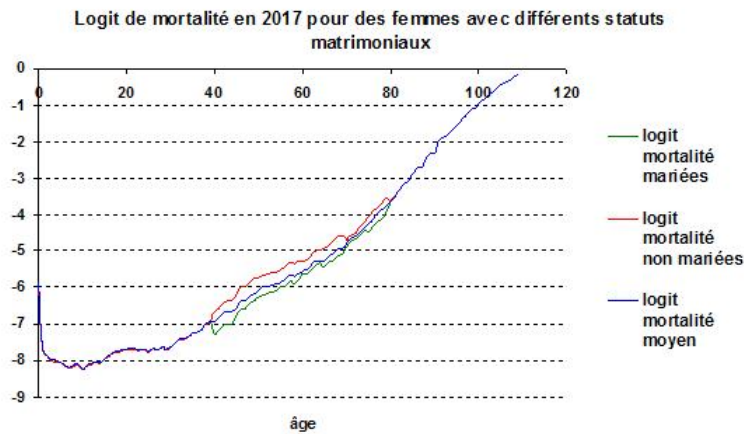


FIGURE 5.44 – Courbes de logit de mortalité des femmes en 2017 pour différents statuts maritaux et par âge

Malgré les remarques concernant les problèmes liés à la variable du statut matrimonial, il apparaît nettement que le statut de marié est protecteur en comparaison des autres statuts maritaux. On remarque aussi que le mariage conserve son statut protecteur au même niveau dans le temps. En revanche, il y a des spécificités liées à l'évolution du statut "non

marié” : aux âges jeunes (avant 50 ans), l’écart de mortalité entre la catégorie ”non marié” et la catégorie moyenne a tendance à diminuer alors qu’aux âges plus élevés (après 50 ans), cet écart a tendance à augmenter.

Aux âges très jeunes, la différence est difficile à prendre en compte car entre 20 et 30 ans par exemple, les individus viennent de se marier et il n’y a pas de raisons claires pour lesquelles il devrait y avoir une influence du statut matrimonial.

De 30 à 50 ans, l’impact protecteur du mariage commence à apparaître mais il a tendance à diminuer dans le temps étant donné le changement de mœurs avec des personnes qui se marient de plus en plus tard ou qui sont en couples sans se marier.

Cependant, à partir de 50 ans, les différences de mortalité sont nettes et se maintiennent jusqu’aux âges élevés (jusque 80 ans) en confrontant le mariage avec le divorce et le veuvage. Comme le niveau de diplôme, le statut matrimonial affecte de manière quasi symétrique l’homme et la femme à la différence de la catégorie professionnelle qui affecte bien plus les hommes que les femmes.

Bilan et discussions

Ces différents tests encouragent à tenir compte de ces caractéristiques dans l’étude la mortalité. Les caractéristiques ont des influences spécifiques dépendant de l’âge et du sexe et il est intéressant de choisir la caractéristique explicative selon l’âge et le sexe de l’individu analysé. Par exemple, si nous cherchons à estimer la mortalité d’une femme, il faudra utiliser le niveau de diplôme et/ou le statut matrimonial plutôt que la catégorie sociale. Ces écarts de mortalité expliquent le fait que les assurances appliquent des règles différentes pour les rentes suivant les caractéristiques des individus.

Variations sur les tranches d’âge

Nous avons travaillé sur des tranches d’âges de 5 ans et nous pouvons observer des phénomènes différents selon la tranche d’âge analysée. Les variations que l’on observe proviennent de deux causes. D’une part, il peut y avoir des variations de précision de données avec l’âge et d’autre part l’impact de la caractéristique dépend de l’âge que l’on considère. Nous aurions pu prendre des tranches d’âge de 10 ans mais les tranches d’âge ne doivent pas être trop larges afin de conserver une indépendance entre les traits et l’âge au sein d’une même tranche d’âge. Pour améliorer ces simulations, il faudrait trouver un moyen de lisser la mortalité sur l’ensemble des tranches d’âges pour avoir des courbes plus régulières.

De plus, l’étude se limite aux âges compris entre 30 ans et 85 ans dans la mesure où l’influence des caractéristiques est petite aux âges faibles et aux âges élevés. Il est difficile

d'avoir un nombre suffisant de données de décès pour ces tranches d'âge. Nous pensons que les courbes de mortalité doivent être proches de la mortalité moyenne pour les âges supérieurs à 85 ans. Par conséquent, nous avons effectué une interpolation linéaire afin de recoller aux courbes nationales de mortalité.

Prise en compte de plusieurs variables en même temps

Nous avons étudié l'influence de différentes caractéristiques indépendamment (CSP, niveau de diplôme et statut matrimonial). L'idéal serait de considérer l'analyse d'un ensemble de caractéristiques en décrivant l'influence jointe de ces caractéristiques. Par exemple, avec deux classes pour un trait x_1 et deux classes pour un trait x_2 ce qui fournit 4 classes de traits possibles. Nous présentons le modèle joint qui décrit la mortalité en fonction de l'âge, et des traits x_1 et x_2 . Ce modèle s'écrit donc sous la forme :

$$\begin{aligned} \forall a \in A_l, \text{ logit}(q(x_1, x_2, a, t)) = & A_1(t) + aA_2(t) + \beta_{1,1}^l(t)1_{\{x_1=x_1^1\}}1_{\{x_2=x_2^1\}} \\ & + \beta_{1,2}^l(t)1_{\{x_1=x_1^1\}}1_{\{x_2=x_2^2\}} + \beta_{2,1}^l(t)1_{\{x_1=x_1^2\}}1_{\{x_2=x_2^1\}} \\ & + \beta_{2,2}^l(t)1_{\{x_1=x_1^2\}}1_{\{x_2=x_2^2\}}, \end{aligned}$$

où les facteurs A_1 et A_2 sont les facteurs du modèle classique de Cairns calibrés sur les tables de mortalité par âge et les facteurs $\beta_{i,j}^l(t)$ sont les facteurs par traits, indépendants des facteurs par âge pour chaque tranche d'âge et qui sont calibrés sur les données de l'EDP.

5.5 Conclusion

Dans ce chapitre, nous avons étudié le modèle de Cairns qui est un modèle standard de mortalité. Ensuite, nous avons étudié qualitativement les déterminants de mortalité en faisant référence à des études économiques. Ces études suggèrent le besoin d'un modèle de mortalité par âge et par traits qu'il faut calibrer sur des données spécifiques. Nous avons introduit un modèle de mortalité par âge et par traits que nous avons calibré sur les données de l'EDP. Ce modèle permet de quantifier l'influence de différentes caractéristiques et nous avons quantifié l'influence spécifique de la catégorie sociale, du niveau d'éducation et du statut matrimonial.

L'objectif de cette étude n'est pas de fournir des tables de mortalité de référence mais plutôt de donner des méthodes d'étude de mortalité par âge et par traits et de voir quels résultats il est possible de déduire. Il faudrait d'ailleurs avoir plus de données afin d'estimer au mieux les différents paramètres du modèle. Il existe de nombreuses pistes d'amélioration avec en particulier une étude statistique rigoureuse pour l'estimation des taux de mortalité et des tendances d'évolution.

Bibliographie

- [1] Antony L. and B. Davalos (2004), “A stochastic model for pricing longevity-linked guarantees”, *Life & Pensions*, 35, 113-136.
- [2] Bensusan H. and N. El Karoui (2009), “Microscopics Models for Populations Trait and Age Structures and Applications for Longevity problem”, Preprint.
- [3] Blake D., A. Cairns and K. Dowd (2006), “Living with mortality : Longevity bonds and other mortality-linked securities”, *British Actuarial Journal*, 12, 153-228.
- [4] Blake D., A. Cairns and K. Dowd (2006), “A two-factor model for stochastic mortality with parameter uncertainty : Theory and calibration”, *Journal of Risk and Insurance*, 73, 687-718
- [5] Blake D., A. Cairns, K. Dowd, G.D. Coughlan, D. Epstein, A. Ong and I. Balevich (2007), “A quantitative comparison of stochastic mortality models using data from England & Wales and the United States”, Working paper, Heriot-Watt University, and Pensions Institute Discussion Paper PI-0701.
- [6] Brouhns N., M. Denuit and J.K. Vermunt (2002), “A Poisson log-bilinear regression approach to the construction of projected life tables”, *Insurance : Mathematics and Economics*, 31, 373-393.
- [7] Cambois E., C. Laborde and J.M Robine (2008), “La « double peine » des ouvriers : plus d’années d’incapacité au sein d’une vie plus courte”, *Populations & Sociétés*, 441, Janvier.
- [8] Devaux M., F. Jusot, A. Trannoy and S. Tubeuf (2007), “Inégalités des chances en santé : influence de la profession et de l’état de santé des parents”, *Questions d’économie de la santé*, IRDES, 118, Février
- [9] Dahl, M. (2004), “Stochastic mortality in life insurance : Market reserves and mortality-linked insurance contracts”, *Insurance : Mathematics and Economics*, 35, 113-136.
- [10] Deaton A. (2001), “Health, Inequality and Economic Development”, NBER WP 8318.
- [11] De Jong P. and L. Tickle (2006), “Extending the Lee-Carter model of mortality projection”, *Mathematical Population Studies*, 13, 1-18.
- [12] Denuit M. and A. Delwaerde (2005), “Construction de tables de mortalité périodiques et prospectives”.

BIBLIOGRAPHIE

- [13] Devineau L. (2009), “La modélisation du risque de mortalité”, *Working Paper*, [http ://fr.milliman.com/publications/bulletin/index.php](http://fr.milliman.com/publications/bulletin/index.php).
- [14] El Karoui N., S. Loisel and C.Mazza (2010), “Detection of changes in longevity trends”, *Working Paper*.
- [15] Gourieroux C. (2008), “Quadratic stochastic intensity and prospective mortality tables Insurance”, *Mathematics and Economics*, 43, 1, 174-184.
- [16] Insee, Direction des statistiques démographiques et sociales (2003), “Echantillon Démographique Permanent : Manuel de l'utilisateur en statistiques et études ”, Intern Paper, [http ://siecles.insee.fr](http://siecles.insee.fr).
- [17] Jusot F. (2004), “Mortalite, revenu et inégalités de revenu en France : une analyse multi-niveaux”, IRDES, Reference BDSP : 309602.
- [18] Jusot F. (2006), “The Shape of the Relationship Between Mortality and Income in France ”, *Annales d'économie et de statistique*, 83-84.
- [19] Lee R.D. and L.R. Carter (1992), “Modeling and forecasting U.S. mortality”, *Journal of the American Statistical Association*, 87, 659-675.
- [20] Lee R.D. and T. Miller (2001), “Evaluating the performance of the Lee-Carter model for forecasting mortality”, *Demography*, 38, 537-549.
- [21] Mejer L. and I. Robert-Bobée (2003), “Mortalité des femmes et environnement familial : Rôle protecteur de la vie de famille”, *Insee Premiere*, 892, Avril.
- [22] Olivieri A. (2001), “Uncertainty in mortality projections : an actuarial perspective Insurance”, *Mathematics and Economics*, 29, 2, 231-245.
- [23] Pla A. and C. Beaumel (2010), “Bilan Démographique 2009 : Deux pacs pour trois mariage”, *Insee Première*, 1276, Janvier.
- [24] Preston S.H. (1975), “The changing relation between mortality and level of economic development”, *Population Studies*, 29 : 231-48.
- [25] Putnam R.D. (1995), “Bowling alone : America's declining social capital”, *Journal of Democracy*, 6 : 65-78.
- [26] Workshop of La Chaire Transition Démographique (2010), “Demographic perspectives : How to measure and deal with uncertainty”, [http ://www.dauphine.fr/fileadmin/mediatheque/chaieres/chairetransition/pdf/Rapport2009.pdf](http://www.dauphine.fr/fileadmin/mediatheque/chaieres/chairetransition/pdf/Rapport2009.pdf).
- [27] Renshaw A.E. and S. Haberman (2003), “Lee-Carter mortality forecasting with age-specific enhancement”, *Insurance : Mathematics and Economics*, 33, 255-272.
- [28] Renshaw A.E. and S. Haberman (2006), “A cohort-based extension to the Lee-Carter model for mortality reduction factors”, *Insurance : Mathematics and Economics*, 38, 556-570.
- [29] Richards S.J., J.G. Kirby and I.D. Currie (2006), “The importance of year of birth in two-dimensional mortality data”, *British Actuarial Journal*, 12, 5-61.

- [30] Robert-Bobée I. and O. Cadot (2007), “Mortalité aux grands âges : encore des écarts selon le diplôme et la catégorie sociale”, *Insee Première*, 112, Février.
- [31] Robert-Bobée I. and C. Monteil (2006), “Différentiels sociaux et familiaux de mortalité aux âges actifs : quelles différences entre les femmes et les hommes”, *Economie et statistique*, 398-399.
- [32] Rodgers G.B. (1979), “Income and inequality as determinants of mortality : an international cross-section analysis”, *Population Studies*, 33 : 343-51.
- [33] Schrager D.F. (2006), “Affine stochastic mortality”, *Insurance : Mathematics and Economics*, 38, 81-97.
- [34] Subramaniam S.V., Blakely T., Kawachi I. (2003), “Income Inequality as a Public Health Concern : Where do we stand?”, *Health Services Research*, 38, 1 : 153-167.
- [35] Vallin J. and C. Lévy (1981), “La mortalité par catégorie socio-professionnelle. Un essai de calcul direct”, *Population*, 36, 4, 938-945.
- [36] Wilkinson R.G. (1992), “Income distribution and life expectancy”, *British Medical Journal*, 304 : 165-68.
- [37] Willets R.C. (2004), “The cohort effect : Insights and explanations”, *British Actuarial Journal*, 10, 833-877.

BIBLIOGRAPHIE

Chapitre 6

Modélisation microscopique de la dynamique de population

Introduction

Ce chapitre aborde la modélisation microscopique de la dynamique d'une population structurée par âge et par traits. Le modèle doit prendre en compte différents événements à l'échelle de l'individu comme les naissances, les décès, l'immigration et les changements de statut, dont les intensités seront des processus stochastiques dépendant des facteurs d'environnement et des caractéristiques des individus considérés.

Motivations

Nous présentons les motivations pratiques de ce type de modélisation dont les applications seront détaillées dans les chapitres suivants de la thèse.

La première motivation est de faire une étude démographique : la situation actuelle invite à modéliser au mieux la dynamique de la population d'un pays. En effet, le débat concernant la longévité des individus est très controversé dans la mesure où les estimations données par les démographes se révèlent souvent assez éloignées de ce qui se passe en réalité. D'un point de vue probabiliste, l'évolution de la longévité n'est pas déterministe mais aléatoire et difficile à prédire. Il reste d'ailleurs de nombreuses questions ouvertes comme le fait de savoir si la longévité est indéfiniment élastique ou s'il y a un âge critique qu'un être humain ne pourra pas dépasser.

Cette motivation démographique est aussi fortement liée à des enjeux économiques puisque l'amélioration de la longévité des personnes âgées a un impact important sur le produit intérieur brut (PIB) des pays ainsi que sur la politique des gouvernements. Le vieillissement de la population est dû notamment à un taux de fécondité actuel faible, et à une amélioration nette des conditions de vie. L'actualité montre que les gouvernements sont soumis au

problème du financement des retraites pour lequel ils doivent adopter des stratégies. Nous proposons un modèle dynamique afin de décrire l'évolution de la population d'un pays. Le modèle pourra être utilisé dans le but de fournir des stratégies qui répondent mieux aux attentes des gouvernements, comme par exemple une immigration optimale ou une réforme sur l'âge de départ à la retraite.

La deuxième motivation est purement financière avec l'étude du risque de longévité. Les compagnies d'assurance ou de réassurance, les fonds de pension, les états sont largement exposés aux risques financiers liés à la longévité des individus qu'ils assurent et à son évolution dans les années à venir. Par conséquent, ces acteurs cherchent à couvrir ce risque à long terme, et ils ont besoin d'un modèle précis et efficace décrivant l'évolution d'une population. Le fait de considérer un modèle microscopique par "traits" permet de tenir compte des caractéristiques individuelles. Cela sera très utile dans l'étude du risque de base d'un portefeuille de rentes d'une compagnie d'assurance : ce risque pourra être quantifié en tenant compte de la mortalité spécifique des individus assurés tout en ayant estimé les paramètres sur la population nationale pour laquelle les données sont plus nombreuses.

Afin de caractériser au mieux les individus de la population, et de les prendre en compte dans la modélisation, nous nous basons sur des études démographiques récentes, et nous nous inspirons largement des travaux mathématiques novateurs développés dans le domaine de la biologie.

Introduction aux modèles de dynamique de population

Dans le domaine de la biologie et de l'écologie, les chercheurs ont proposé des modèles de dynamique de population afin de comprendre l'évolution d'une population et de prédire son comportement dans le futur.

Il existe deux types de modèles : les modèles macroscopiques qui décrivent la structure globale d'une population et les modèles microscopiques qui décrivent précisément chaque individu de la population. La modélisation macroscopique, la plus classique, est généralement déterministe et conduit à la résolution d'équations aux dérivées partielles. La modélisation microscopique est plus récente et plus complexe du fait de la stochasticité qui lui est intrinsèque. Les deux types de modélisation présentent des avantages et des inconvénients [21].

Dans l'étude macroscopique de l'évolution d'une population structurée par âge, il existe de nombreux modèles déterministes d'évolution comme ceux de Sharpe et Lotka [37], McKendrick [31], Von Foerster [41], Perthame [34] qui décrivent l'évolution de la densité de la population. Par exemple, l'équation de McKendrick-Von Foerster qui traduit l'évolution

dans le temps de la densité g d'une population structurée par âge est de la forme :

$$\left(\frac{\partial g}{\partial t} + \frac{\partial g}{\partial a}\right)(a, t) = -d(a)g(a, t), \quad g(0, t) = \int_0^\infty b(a)g(a, t)da. \quad (6.1)$$

Cette équation traduit le fait que la densité par âge de la population évolue dans le temps par l'intermédiaire des intensités de mortalité et de natalité. L'intensité de mortalité par âge $d(a)$ apparaît dans l'évolution de cette densité du fait de la prise en compte des décès des individus et le taux de natalité par âge $b(a)$ intervient dans la condition au bord à l'âge 0 en ajoutant les naissances d'individus d'âge 0 : le nombre de naissance à un instant t dépend bien évidemment de la structure de la population à cette même date. Ces modèles macroscopiques, basés sur des équations aux dérivées partielles, permettent d'obtenir des scénarios moyens d'évolution d'une population. Une description des modèles les plus classiques est présentée dans [43]. Les modèles macroscopiques les plus récents étudient des populations structurées par âge et par traits comme ceux de Ernande [15], Rotenberg [36], Perthame et Ryzhik [34], et Mischler [33].

Les modèles microscopiques se proposent de modéliser des scénarios "aléatoires" d'évolution de population. Ce point de vue n'est pas nouveau puisque toute une branche des probabilités s'est développée à partir du processus de branchement introduit par Galton et Watson en 1873 afin de modéliser l'extinction des familles nobles en Angleterre. Dès lors, la théorie s'est enrichie dans de nombreuses directions. Des processus de branchement pour des populations structurées par âge qui généralisent le processus de Galton Watson [19] ont été étudiés par Bellman et Harris [4] et par Athreya et Ney [3]. De plus, Kendall [25], Crump et Mode [11, 12], Jagers [22, 23], Doney [14] se sont intéressés à l'étude de processus de naissance et mort où chaque individu peut donner naissance à un enfant à différentes dates aléatoires durant sa vie, le taux de naissance pouvant dépendre de l'âge. Des phénomènes d'interactions entre les individus ont été introduits, influençant soit le taux de natalité [42, 38] soit le taux de mortalité [35, 22].

Une approche récente introduite par N. Fournier et S. Méléard [18] propose une évolution microscopique aléatoire d'une population structurée par traits. Ce modèle s'inspire des travaux de Bolker-Pacala [6], Dieckmann-Law [15] qui a donné naissance au processus appelé BPLD qui sera à la base de notre étude. Ces travaux ont été poursuivis par de nombreux chercheurs [9, 26] qui en proposent des améliorations.

V.C. Tran [39, 40, 30, 17] a adapté ce point de vue à la modélisation d'une population structurée par âge et par traits, ce qui fournit un cadre bien adapté à l'étude de l'évolution des populations humaines. Son objectif est de permettre une étude qualitative de phénomènes écologiques d'évolution.

Les modèles microscopiques sont par nature plus complexes puisqu'ils modélisent la population à l'échelle de l'individu et qu'ils sont structurellement aléatoires. La considération de l'évolution d'un échantillon de la population (lorsque celle ci est très grande) nécessite un

effort supplémentaire de modélisation afin de décrire l'évolution de chaque individu de la population qui est caractérisé par son âge et par des *traits*. Ces traits sont des caractéristiques individuelles correspondant à des caractères morphologiques, physiologiques, génétiques ou comportementaux. La modélisation microscopique fournit des scénarios d'évolution dont on peut extraire une information de sortie plus riche que celle provenant des modèles macroscopiques.

Présentation du modèle

Dans ce chapitre, nous présentons une approche dynamique de type micro-macro de la modélisation de l'évolution d'une population humaine, en partant de l'échelle de l'individu. Lorsque la taille de la population est grande, le modèle microscopique peut être approximé par un modèle macroscopique généralisant des modèles macroscopiques d'évolution. Nous cherchons à prendre en compte un nombre important d'événements influant sur l'évolution de la population, comme les naissances, les décès, mais aussi l'immigration et l'évolution de caractéristiques d'individus dans le temps. Des études récentes ont montré que la mortalité humaine est influencée par des facteurs stochastiques exogènes [7, 28]. Nous proposons un moyen de prendre en compte ces considérations. Les intensités d'évolution peuvent dépendre du temps et être stochastiques, les traits des individus évoluer dans le temps de manière aléatoire et l'immigration prise en compte.

Ce modèle microscopique permet d'obtenir des informations à l'échelle macroscopique, ce qui permet de vérifier la cohérence du modèle. Dans le cadre des modélisation de S. Méléard [18] et V.C. Tran [39], il a été démontré que lorsque la taille de la population est grande, le processus microscopique de naissance et de mort converge vers un processus à valeurs mesures dont la densité suit une équation qui généralise les équations classiques de McKendrick-Von Foerster [31, 41, 34]. Inspirés par ces développements, nous proposons une approximation macroscopique de notre modèle avec une extension stochastique de ces équations macroscopiques. L'étude fournit des résultats macroscopiques très intéressants.

6.1 Introduction aux processus d'évolution de population

Dans cette partie, nous rappelons la modélisation des processus de naissance et mort, qui sont des processus de Markov dont l'évolution est constante entre deux temps de saut aléatoires. Ces événements représentent des instants de naissance ou de mort d'individus. Les processus de Poisson ponctuels jouent un rôle très important dans cette étude. Nous en rappelons les principales propriétés. Pour plus d'informations, le lecteur pourra se référer à [32].

6.1.1 Processus ponctuels et processus de Poisson

Définition 6.1. *Un processus ponctuel sur \mathbb{R}_+ est décrit par la donnée d'une suite croissante de temps aléatoires, définis sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$,*

$$0 < T_1 < T_2 < \dots < T_n < \dots \quad p.s.$$

tels que la suite $T_n \rightarrow +\infty$ presque sûrement lorsque n tend vers l'infini.

La fonction aléatoire de comptage $(N_t)_{t \geq 0}$ associée au processus ponctuel $\{T_n, n \in \mathbb{N}\}$ est définie par :

$$N_t = \sup\{n, T_n \leq t\} = \sum_{j \geq 1} 1_{T_j \leq t} = \sum_{j \geq 1} 1_{(0, t]}(T_j)$$

La définition de la suite T_n suggère la définition des temps d'attente τ_n entre deux événements successifs :

$$\tau_1 = T_1, \quad \tau_2 = T_2 - T_1, \quad \tau_n = T_n - T_{n-1}, \dots$$

Le processus N_t correspond au nombre d'événements qui se sont produits avant l'instant t et $N_t - N_s$ correspond au nombre d'événements ayant eu lieu pendant l'intervalle $]s, t]$. Par définition, le processus N_t est continu à droite.

Définition 6.2. *On dit que le processus (N_t) ou que le processus ponctuel (T_n) sous-jacent est un processus de Poisson d'intensité λ si le processus $(N_t, t \geq 0)$ est à accroissements indépendants et stationnaires, c'est à dire*

1. *Pour tous $t_0 < t_1 < \dots < t_n$ dans \mathbb{R}_+ , les accroissements $(N_{t_j} - N_{t_{j-1}}, 1 \leq j \leq n)$ sont des variables aléatoires indépendantes.*
2. *Pour $0 \leq s < t$, la loi de $N_t - N_s$ ne dépend de s et de t que par la différence $t - s$. Elle est donc égale à la loi de N_{t-s} . C'est la propriété de stationnarité des accroissements.*
3. *La loi de N_t est la loi de Poisson de paramètre λt*

Il est connu que l'hypothèse d'accroissements indépendants et stationnaires suffit à prouver l'existence d'un paramètre $\lambda > 0$ tel que pour tous $0 < s < t$, $N_t - N_s$ soit une variable aléatoire de Poisson de paramètre $\lambda(t - s)$. Cette propriété permet de justifier l'appellation de "processus de Poisson". Ce paramètre λ est appelé intensité du processus de Poisson et représente le nombre moyen d'événements qui se produisent pendant une unité de temps. En effet, cela provient du fait que

$$\mathbb{E}(N_{t+s} - N_t) = \lambda s.$$

On démontre aussi que les variables τ_i suivent des lois exponentielles indépendantes de paramètre λ . Cette propriété est caractéristique et aurait pu être prise comme définition d'un processus ponctuel de Poisson.

La théorie des martingales conduit à une autre caractérisation du processus de Poisson. Introduisons la filtration canonique (\mathcal{F}_t^N) engendrée par le processus, complétée et rendue continue à droite. Il est facile de montrer que le processus de Poisson compensé $\tilde{N}_t = N_t - \lambda t$ est une martingale continue à droite et limitée à gauche (càd-làg), de variation quadratique $\langle \tilde{N}_t \rangle = \lambda t$. Il est un peu plus délicat de montrer la réciproque, à savoir vérifier que ces deux propriétés suffisent à montrer que le processus de comptage (N_t) est un processus de Poisson.

6.1.2 Processus de naissance et mort

Les processus de Poisson peuvent être utilisés dans l'étude de dynamique de population afin de modéliser l'occurrence des naissances et des décès. Nous développons certains aspects qui seront utiles dans la suite.

Exemple simple de processus de naissance et mort

Considérons le cas simple de l'évolution d'une population dans laquelle il se produit uniquement des naissances, dont le nombre est modélisé par un processus de Poisson (N_t^b) d'intensité b . Cette intensité b représente le nombre moyen de naissances par unité de temps. Utilisons le même type de modélisation pour les décès en introduisant un processus de Poisson N_t^d d'intensité d indépendant de (N_t^b) . Un processus de naissance et mort de la population correspond à la prise en compte simultanée des événements de naissance et de mort, c'est à dire le processus $Z_t = Z_0 + N_t^b - N_t^d$, arrêté au premier instant où il devient négatif.

Tout se passe comme si le processus lançait deux "horloges" B et D, indépendantes et exponentielles aux taux respectifs b et d . Si c'est l'horloge B qui sonne en premier, il y a une naissance alors que si c'est l'horloge D, il y a un décès.

Le processus de Poisson $N_t^{b,d} = N_t^b + N_t^d$, décrivant l'ensemble des événements de naissance et de mort, a pour intensité $b + d$. Le processus Z peut être réinterprété comme suit :

- Z_0 correspond à la taille initiale de la population qui suit une loi prédéfinie.
- À chaque temps de saut T_n de $N^{b,d}$, il peut se produire une naissance (saut de +1) ou un décès (saut de -1) avec les probabilités respectives $\frac{b}{b+d}$ et $\frac{d}{b+d}$. La détermination de l'événement est associée à la réalisation d'une variable de Bernoulli Y_n .
- Les variables aléatoires T_n ont des accroissements $\tau_n = T_n - T_{n-1}$ indépendants et de loi exponentielle de paramètre $b + d$; les variables de Bernoulli Y_n sont elles aussi indépendantes et indépendantes des T_n .

Tant que $Z_t > 0$, le processus Z vérifie :

$$Z_t = Z_0 + \sum_{n=1}^{N_t^{b,d}} Y_n = Z_0 + \sum_{n \geq 1} 1_{\{T_n \leq t\}} Y_n \quad (6.2)$$

Remarque 6.1. Dans la suite nous aurons besoin de modéliser le fait que dans certaines situations, certains événements n'induisent pas de modification de la taille de la population mais plutôt de sa structure. Nous les prendrons en compte en associant un troisième état $Y_i = 0$ aux variables Y_i ainsi qu'une probabilité de transition associée.

Processus de naissance et de mort et mesure de Poisson ponctuelle Cette simple formulation peut être étendue à des situations plus générales. Par exemple, nous nous intéressons au processus $Z_t(f)$ définie pour toute fonction mesurable f bornée par :

$$Z_t(f) = Z_0 + \sum_{n=1}^{N_t^{b,d}} f(Y_n)$$

Prendre $f(y) = 1_{\{y=1\}}$ permet notamment de considérer des notions relatives aux naissances uniquement alors que prendre f égale à l'identité permet de retrouver le processus de naissance et de mort. En pratique, la fonction $f(Y_n)$ peut aussi s'écrire comme $f(Y_n) = f(1)1_{\{Y_n=1\}} + f(-1)1_{\{Y_n=-1\}}$, ce qui conduit à la représentation de Z comme

$$Z_t(f) = Z_0 + f(1)N_t^b + f(-1)N_t^d$$

La dynamique du processus $Z_t(f)$ peut être décrite comme celle des processus de Poisson par une propriété de martingale. En effet, le processus $\tilde{Z}_t(f)$ défini par :

$$\tilde{Z}_t(f) = Z_0 + f(1)\tilde{N}_t^b + f(-1)\tilde{N}_t^d$$

est une martingale par rapport à la tribu engendrée par le couple (N_t^b, N_t^d) , de variation quadratique $\langle \tilde{Z}_t(f) \rangle = \int_0^t (bf^2(1) + df^2(-1)) ds$.

La martingale $\tilde{Z}_t(f)$ est la différence des deux processus $Z_t(f)$ et $Z_t^c(f)$ à variation finie où $Z_t^c(f)$ est défini par $Z_t^c(f) = \int_0^t (bf(1) + df(-1)) ds$. Ce dernier est un processus croissant prévisible appelé le *compensateur de $Z_t(f)$* . La quantité $\mu(f) = [bf(1) + df(-1)]$ est appelée *l'intensité* et $\nu(f) = \frac{b}{b+d}f(1) + \frac{d}{b+d}f(-1)$ correspond à *l'intensité normalisée*

Mesures de Poisson ponctuelles L'utilisation du formalisme des mesures est bien adaptée pour décrire ces propriétés. Les mesures de Dirac en un point (s, y) notées $\delta_{(s,y)}$, dont l'intégrale par rapport à f vaut $\delta_{(s,y)}(f) = f(s, y)$, jouent un rôle fondamental. Dans les développements précédents, nous n'avons pas vraiment utilisé le fait que les variables Y_n ne prennent qu'un nombre fini de valeurs. Ainsi, nous supposons désormais quelles prennent leurs valeurs dans un espace mesurable (E, \mathcal{E}) , sur lequel est définie une mesure μ de masse finie. Cet ensemble est appelé *l'espace des marques*

1. Soit $(T_n)_{n \geq 0}$ les temps de sauts d'un processus de Poisson sur \mathbb{R}^+ d'intensité $\mu(E)$ et $(Y_n)_{n \geq 0}$ une suite de variables à valeurs dans E indépendantes et de loi $\nu(dy) =$

$\mu(dy)/\mu(E)$, indépendante des T_n . La mesure ponctuelle

$$Z(dt, dy) = \sum_{n \geq 1} \delta_{(T_n, Y_n)}(dt, dy)$$

à laquelle nous associons la mesure compensée $Z^c(dt, dy) = dt\mu(dy) = dt\mu(E)\nu(dy)$ caractérisée par le fait que pour toute fonction g telle que $\int_0^{+\infty} ds \int_E |g(s, y)|\mu(dy) < \infty$ le processus à variation finie

$$\tilde{Z}_t(g) = \int_0^t \int_E g(s, y) Z(ds, dy) - \int_0^t ds \int_E g(s, y) \mu(dy)$$

est une martingale. Si de plus $\int_0^{+\infty} ds \int_E |g^2(s, y)|\mu(dy) < \infty$, la variation quadratique de $\tilde{Z}_t(g)$ est égale à $\langle \tilde{Z}_t(g) \rangle = \int_0^t ds \int_E g^2(s, y) \mu(dy)$ correspondant à la définition précédente dans un cadre plus général

En particulier, pour tout $A \subset E$, le processus $Z_t(1_A)$ est un processus de Poisson de paramètre $\mu(A)$ et la mesure s'appelle *mesure ponctuelle de Poisson d'intensité μ* .

2. Cette notion peut être étendue aux mesures σ -finies de la manière suivante. Soit μ^n une suite de mesures de masses finies, à supports disjoints, dont la somme dénombrable est la mesure μ . À la suite de mesures μ^n , nous associons une suite de processus de Poisson marqués Z^n d'intensité μ^n indépendants les uns des autres. Le processus Z correspond à un processus de Poisson marqué d'intensité μ , dont nous ne pouvons plus décrire directement comment le générer.

Processus ponctuels discrets Les mesures ponctuelles peuvent être plus générales que les mesures ponctuelles de Poisson. En particulier, il est possible d'affaiblir les hypothèses d'indépendance et de stationnarité des marques par rapport aux temps de sauts auxquels se produisent les événements. Le prix à payer est bien sûr l'abandon du caractère poissonnien, au profit des propriétés de martingales.

- Nous supposons maintenant qu'une filtration (\mathcal{F}_t) est donnée par rapport à laquelle toutes les quantités d'intérêt seront adaptées et considérons une mesure d'intensité de la forme

$$\mu(ds, dy) = \lambda(s) ds \nu(s, dy)$$

où $\lambda(\cdot)$ est un processus prévisible positif, et $\nu(\cdot, dy)$ est un processus aléatoire prévisible à valeurs dans l'espace des probabilités sur E . Autrement dit, la mesure ν garde la même interprétation que dans les exemples précédents : lorsqu'une cloche sonne entre s et $s + ds$ signalant l'occurrence d'un événement, la marque est tirée au hasard suivant la loi $\nu(s, dy)$.

- Soit $(T_n)_{n \geq 0}$ un processus ponctuel à valeurs dans \mathbb{R}_+ (au sens de la première définition) tel que les v.a. T_n soient des temps d'arrêt de la filtration (\mathcal{F}_t) . Soit d'autre

part, $(Y_n)_{n \geq 0}$ une suite de variables aléatoires \mathcal{F}_{T_n} – mesurables, prenant leurs valeurs dans E .

Le processus ponctuel $Z_t(ds, dy) = \sum_{n \geq 1} \delta_{(T_n, Y_n)}(ds, dy) 1_{\{T_n \leq t\}}$, à valeurs dans l'ensemble des mesures discrètes, est adapté et constant entre deux temps de sauts. C'est la raison pour laquelle nous l'appellerons *processus ponctuel discret*.

- Nous dirons que $\mu(s, dy) = \lambda(s) \nu(s, dy)$ est la *mesure d'intensité* de Z si pour tout processus prévisible g positif tel que $\mathbb{E}(\int_{\mathbb{R}_+} ds \int_E g(s, y) \mu(s, dy)) < +\infty$, le processus à variation finie,

$$\tilde{Z}_t(g) = \sum_{T_n \leq t} g(T_n, Y_n) - \int_0^t ds \int_E g(s, y) \mu(s, dy)$$

est une martingale. Lorsque la mesure $\mu(s, \cdot)$ est seulement σ -finie, les mêmes définitions restent valides.

- La notion de martingale associée au processus $\tilde{Z}_t(g)$ reste valide pour des processus prévisibles tels que $\mathbb{E}(\int_{\mathbb{R}_+} ds \int_E g(s, y)^2 \mu(s, dy)) < +\infty$, par passage à la limite ; une telle martingale peut ne pas être à variation finie.

Liens avec les mesures de Poisson Nous nous intéressons plus particulièrement aux processus ponctuels discrets dont la mesure d'intensité sur E , $\mu(s, dy)$, admet une densité prévisible $k(s, y)$ par rapport à une mesure (éventuellement σ -finie) de référence non aléatoire $P(dy)$. Il est possible de simuler cette intensité par une méthode de type rejet en introduisant une composante supplémentaire θ à valeurs dans \mathbb{R}_+ muni de la mesure de Lebesgue et en remarquant que,

$$\mu(s, dy) = \int_{\mathbb{R}_+} 1_{\{\theta \leq k(s, y)\}} P(dy) d\theta.$$

Considérons une mesure aléatoire de Poisson $Q(dt, dy, d\theta) = \sum_{n \geq 1} \delta_{(T_n, Y_n, \Theta_n)}(dt, dy, d\theta)$, adaptée à la filtration (\mathcal{F}_t) , d'intensité $P(dy) d\theta$ et dont les marques sont à valeurs dans $E \times \mathbb{R}_+$. En intégrant sur la composante θ , le processus ponctuel discret

$$Z_t(ds, dy) = \int_{\mathbb{R}_+} 1_{\{\theta \leq k(s, y)\}} Q_t(ds, dy, d\theta) = \sum_{T_n \leq t} \delta_{(T_n, Y_n)}(ds, dy) 1_{\{\Theta_n \leq k(T_n, Y_n)\}}$$

admet pour intensité la mesure prévisible $\mu(s, dy)$.

Remarque 6.2. Contrairement à ce qui se passe pour la mesure intensité $\mu(s, dy)$, le processus ponctuel $Z_t(ds, dy)$, induit par $Q(dt, dy, d\theta)$, ne peut pas être défini sans faire référence à la suite des marques Θ_n car la mesure $\sum_{n \geq 1} \delta_{(T_n, Y_n, \Theta_n)}$ n'est pas une mesure produit. En effet, le tirage de Y_n et de Θ_n se fait conditionnellement au tirage de la variable T_n .

6.1.3 Processus ponctuel discret d'évolution

L'étude d'un modèle de dynamique de population microscopique est plus complexe que le cadre décrit précédemment. En effet, il est nécessaire de décrire une information détaillée par individu afin de modéliser les évolutions de structure de population. Par ailleurs, ces individus sont caractérisés par des traits x (à valeurs dans un espace χ) qui ne changent que ponctuellement dans le temps. Ainsi, nous prenons en compte en plus le phénomène d'évolution des traits des individus par l'ajout d'un état e . La prise en compte de l'âge sera faite dans la dernière étape de la modélisation en Section 6.3.

Le processus d'évolution de la population prend ses valeurs dans l'espace des mesures ponctuelles finies sur χ qui sont de la forme $z(dx) = \sum_{j \geq 1} \delta_{x_j}(dx)$. Le nombre d'individus dans la population à l'instant t est noté $N_t = \langle z_t, 1 \rangle$.

Décrivons le mécanisme d'évolution de la population d'un état à un autre : partant d'une population Z_{n-1} de taille N_{n-1} , un individu est tiré au hasard : ses traits déterminent avec quelle probabilité il va se reproduire, mourir, changer de traits, ou ne rien faire. Les événements d'évolution sont modélisés par l'intermédiaire d'une variable U_n prenant des valeurs discrètes $\epsilon \in \{b, d, e\}$. Ensuite seulement, les traits du "nouvel" individu sont tirés suivant une loi dépendant du passé, de la variable U_n et des traits de l'individu sélectionné. La taille de la population évolue comme un processus de naissance et de mort à trois états $\{b, d, e\}$ qui n'est pas nécessairement poissonnien.

Le processus d'évolution $Z_t(dx) = \sum_{j=1}^{N_t} \delta_{X_j(t)}(dx)$ qui est la somme de ses sauts, est représenté par un processus ponctuel discret

$$Z_t(dx) = Z_0(dx) + \sum_{T_n \leq t} \left(\sum_{j=1}^{N_n} \delta_{X_j^n}(dx) - \sum_{j=1}^{N_{n-1}} \delta_{X_j^{n-1}}(dx) \right),$$

où les temps d'arrêt T_n sont les temps de sauts effectifs du processus ponctuel discret Z . Désignons par J_{n-1} l'individu tiré dans la population Z_{n-1} de traits $X_{J_{n-1}}^{n-1}$, et associons lui une variable U_n à valeurs dans $\{b, d, e\}$, un nouvel individu I_n de trait $X_{I_n}^n$. Le saut de Z en T_n est $\delta_{X_{I_n}^n}(dx)$ pour une naissance ($U_n = b$), $-\delta_{X_{J_{n-1}}^{n-1}}(dx)$ pour un décès ($U_n = d$), et $(\delta_{X_{J_{n-1}}^n}(dx) - \delta_{X_{J_{n-1}}^{n-1}}(dx))$ pour un changement de trait ($U_n = e$), car dans ce cas $I_n = J_{n-1}$. Le processus $Z_t(dx) - \int_0^t \int_{\chi} Z_s(dx') \phi(s, x', dx) ds$ est une martingale où $\phi(s, x', dx)$ s'écrit sous la forme :

$$\begin{aligned} \phi(s, x', dx) = & \lambda_s [1_{\epsilon=b} \pi^b(s, x') \nu^b(s, x', dx) - 1_{\epsilon=d} \pi^d(s, x') \delta_{x'}(dx) \\ & + 1_{\epsilon=e} \pi^e(s, x') (\nu^e(s, x', dx) - \delta_{x'}(dx))] \end{aligned}$$

avec π^ϵ des probabilités représentant le tirage de la variable ϵ et ν^ϵ des probabilités liées au tirage du trait du nouvel individu. La présence du $Z_s(dx')$ est liée à la sélection de l'individu

au sein de la population.

Pour toutes les valeurs de ϵ , les processus $\lambda_t \pi^\epsilon(t, x')$ sont prévisibles et sont généralement supposés bornés. Pour $\epsilon \neq d$, les probabilités $\nu^\epsilon(t, x', dx)$ sont prévisibles et sont aussi supposées à densité bornée $k^\epsilon(t, x', x)$ par rapport à une mesure $P(dx)$ de référence sur χ . La probabilité $\nu^d(t, x', dx)$ qui n'apparaît pas explicitement dans la formule nous sera utile dans la suite.

Remarque 6.3. ϕ n'est pas forcément une mesure positive alors que son intégrale $\int_\chi Z_s(dx') \phi(s, x', dx)$ l'est nécessairement.

Représentation poissonniene du processus ponctuel discret d'évolution

Puisque la mesure signée ϕ admet une densité par rapport à une mesure non aléatoire, nous pouvons utiliser la construction précédente basée sur les mesures aléatoires de Poisson. Toutefois, il sera utile d'introduire une famille de mesures supplémentaires. Pour chaque valeur de ϵ , nous définissons une mesure $\hat{\mu}(t, x', \epsilon, dx) = \lambda_t \pi^\epsilon(t, x') k^\epsilon(t, x', x) P(dx)$. La densité $m(t, i, X_i, x, \epsilon) = \lambda_t \pi^\epsilon(t, X_i) k^\epsilon(t, X_i, x)$, est construite à partir des variables $((i, X_i); i \leq N_t)$ qui définissent la mesure ponctuelle Z_t .

Afin de représenter la mesure prévisible $\int_\chi Z_t(dx') \hat{\mu}(t, x', \epsilon, dx)$, on procède de la manière suivante. La représentation de cette densité peut se faire en introduisant un paramètre θ et en définissant les fonctions suivantes

$$\begin{aligned} \kappa(t, i, X_i, \theta, x, b) &= 1_{\{0 \leq \theta \leq m_1(t, i, X_i, x)\}}, \\ \kappa(t, i, X_i, \theta, x, d) &= 1_{\{m_1(t, i, X_i, x) \leq \theta \leq m_2(t, i, X_i, x)\}}, \\ \kappa(t, i, X_i, \theta, x, e) &= 1_{\{m_2(t, i, X_i, x) \leq \theta \leq m_3(t, i, X_i, x)\}}, \end{aligned}$$

où les fonctions m_1 , m_2 et m_3 définies par $m_1(t, i, X_i, x) = m(t, i, X_i, x, b)$, $m_2(t, i, X_i, x) = m_1(t, i, X_i, x) + m(t, i, X_i, x, d)$ et $m_3(t, i, X_i, x) = m_2(t, i, X_i, x) + m(t, i, X_i, x, e)$.

Nous considérons ensuite une mesure de référence η sur l'espace $\Xi = \mathbb{N} \times \mathbb{R}_+ \times \chi$, donnée par $\eta(di, d\theta, dx) = n(di) d\theta P(dx)$, où n est la mesure de comptage sur \mathbb{N} , $d\theta$ la mesure de Lebesgue sur \mathbb{R}_+ et P la probabilité de référence sur χ .

Nous introduisons une mesure aléatoire de Poisson Q sur l'espace des marques $\mathbb{N} \times \mathbb{R}_+ \times \chi$ dont la mesure d'intensité est la mesure produit $\eta(di, d\theta, dx') = n(di) d\theta P(dx')$. Désignons par S_n les temps de sauts de la mesure aléatoire de Poisson Q et par (I_n, θ_n, X'_n) les marques associées qui sont indépendantes de S_n . La mesure $\hat{Z}_t(\epsilon, dx)$ définie par $\hat{Z}_t(\epsilon, dx) = \int_{\mathbb{N}} \int_{\mathbb{R}_+} \kappa(t, i, X_i, \theta, x', \epsilon) Q_t(di, d\theta, dx')$ a pour intensité $\int_\chi Z_t(dx') \hat{\mu}(t, x', \epsilon, dx)$ et n'est autre que la mesure

$$\hat{Z}_t(\epsilon, dx) = \sum_{S_n \leq t} 1_{I_n \leq N_t} \kappa(S_n, I_n, X_{I_n}, \Theta_n, X'_n, \epsilon) \delta_{X'_n}(dx)$$

Le processus ponctuel discret d'évolution est construit de manière un peu plus complexe en prenant des intégrales de mesures,

$$\begin{aligned} Z_t(dx) &= Z_0(dx) + \int_0^t \int_{\Xi} 1_{i \leq N_{s-}} K(s, i, X_i, \theta, x', dx) Q(ds, di, d\theta, dx'), \\ \text{où} \quad K(s, i, X_i, \theta, x', dx) &= \kappa(s, i, X_i, \theta, x', b) \delta_{x'}(dx) \\ &\quad - \kappa(s, i, X_i, \theta, x', d) \delta_{X_i}(dx) + \kappa(s, i, X_i, \theta, x', e) [\delta_{x'}(dx) - \delta_{X_i}(dx)] \end{aligned} \quad (6.3)$$

Notons que les sauts du processus $Z_t(dx)$, correspondant aux temps T_n , sont effectivement suivis d'un changement dans la population. Ils sont donc a priori beaucoup moins nombreux que les temps de saut S_n de la mesure de Poisson Q .

Remarque 6.4. *La mesure ponctuelle de Poisson Q n'est pas facile à interpréter puisque son intensité est de masse infinie mais elle reste possible à décrire et à simuler (voir Section 6.1.2). Une manière de mieux comprendre les choses est de supposer que la taille de la population reste en dessous d'un seuil \bar{N} et que la variable θ est majorée par un taux $\bar{\lambda}$. Cela permet d'avoir une intensité finie et de conserver un cadre poissonnien avec des variables exponentielles S_n de paramètre constant $\bar{N}\bar{\lambda}$ indépendant de la structure de la population. Dans la simulation du processus, il sera plus efficace de ne pas supposer la taille de la population majorée quitte à perdre le caractère poissonnien puisque la loi des S_n dépendra de la population Z_n .*

Cette présentation, certes assez complexe, a le grand intérêt de se prêter facilement à diverses généralisations, comme la prise en compte de phénomènes d'immigration ou de dépendance entre les individus. Néanmoins, la principale extension pour la dynamique des individus est la prise en compte de l'âge des individus. Cette généralisation est proposée en Section 6.3 en s'appuyant sur les travaux de V.C. Tran [39]. La principale difficulté provient du fait que le processus d'évolution n'est plus constant par intervalle mais évolue continûment entre les sauts. La présentation ci-dessus sera à la base de la modélisation de ce chapitre où le processus d'évolution sautera à des instants générés par une mesure ponctuelle de Poisson.

6.2 Description des paramètres d'évolution de la population

Nous spécifions maintenant les paramètres nécessaires pour décrire l'évolution microscopique de la population. Il faut tenir compte dès lors du paramètre de l'âge, notamment dans la modélisation des intensités de naissance, de décès, ainsi que des probabilités de changement d'un trait à un autre. Il faut aussi prendre en compte l'impact de l'environnement sur ces différents paramètres.

Les intensités d'évolution permettent de décrire les différents événements affectant l'évolution de la population. Nous allons introduire des facteurs aléatoires exogènes qui vont

permettre de décrire la dynamique de ces intensités dans le temps. Nous formulons les hypothèses sur ces intensités qui sont nécessaires à la définition du processus d'évolution et qui permettent de définir la probabilité de survie. Enfin, nous analysons la différence existante entre une intensité d'évolution et un taux d'évolution.

6.2.1 Intensités individuelles d'évolution

Nous nous intéressons à la description des intensités individuelles d'évolution du processus de dynamique de population. Chacun des événements est décrit afin de comprendre la manière dont la population évolue dans le temps.

Intensité de natalité

Les naissances dépendent des caractéristiques individuelles dont l'âge et certains traits. À chaque naissance, il y a une *mutation des traits* dans la mesure où les descendants n'ont pas forcément les mêmes caractéristiques que leurs parents. Les traits d'un descendant sont tirés dans une distribution de traits qui pourra éventuellement dépendre des traits des parents.

Un individu de trait $x \in \chi$ et d'âge $a \in \mathbb{R}_+$ peut donner naissance à un nouvel individu avec une intensité $b(x, a)$, où b est une fonction de $\tilde{\chi} \times \mathbb{R}$ dans \mathbb{R}_+ . Les nouveaux nés seront caractérisés par un âge égal à 0 et par un trait $x' \in \chi$. Le trait $x' \in \chi$ est la réalisation d'une variable aléatoire de loi $K^b(x, a, dx')$ qui représente la distribution de la descendance.

L'expression de la fonction b permet de prendre en compte différentes phases de vie de l'individu avec une phase de croissance et une phase de reproduction. L'âge à maturité correspond à l'âge de passage de la période juvénile à la période adulte. Nous pouvons aussi tenir compte du fait que la phase de reproduction baisse d'intensité lorsque l'âge de l'individu vieillit.

Intensité de mortalité

La mortalité individuelle dépend elle aussi des caractéristiques de l'individu. Bien qu'il soit possible de modéliser l'interaction qui pourrait exister entre les individus comme cela est fait dans [39, 40, 30, 17], nous ne tiendrons pas compte ici de la mortalité par compétition.

Un individu d'âge $a \in \mathbb{R}_+$ et de trait $x \in \chi$ peut décéder avec une intensité $d(x, a)$. La forme de l'intensité d permet de tenir compte de phénomènes naturels comme une augmentation de la probabilité de décès lorsque l'individu vieillit. De plus, certaines caractéristiques seront déterminantes dans le niveau de mortalité de l'individu au cours de sa vie.

Intensité d'évolution de traits

Concernant l'évolution des traits d'un individu dans le temps, elles sont aussi prises en compte à l'échelle individuelle. Ces évolutions de traits permettent de modifier la structure de la population dans le temps.

Dans cette modélisation, les traits d'un individu peuvent évoluer dans le temps. Un individu d'âge a et de trait x peut changer de trait avec une intensité d'évolution $e(x, a)$ et acquérir un nouveau trait x' . Le trait x' est tiré suivant une distribution de traits $K^e(x, a, x')$ qui dépend des caractéristiques initiales de l'individu. La forme de l'intensité e décrit le taux de passage d'un trait à un autre et dépendra aussi de l'âge de l'individu.

6.2.2 Évolution des intensités individuelles dans le temps

Historiquement, les intensités d'évolution peuvent évoluer dans le temps et de manière aléatoire. Il faut essayer de tenir compte de la dépendance temporelle des intensités d'évolution. La prise en compte de cette évolution peut se faire à l'aide de facteurs exogènes. Ainsi, nous introduisons des facteurs aléatoires associés aux différentes intensités d'évolution.

Nous notons N_b , N_d et N_e les nombres des facteurs aléatoires exogènes associés à la natalité, la mortalité et l'évolution des traits. Les facteurs Y_t^b , Y_t^d et Y_t^e sont à valeurs respectivement dans \mathbb{R}^{N_b} , \mathbb{R}^{N_d} et \mathbb{R}^{N_e} et permettent de tenir compte de l'évolution (éventuellement stochastique) des intensités dans le temps. Ainsi, les intensités dépendent de ces facteurs exogènes et nous nous intéressons plus précisément à l'intensité de natalité $b(x, a, Y_t^b)$, à l'intensité de mortalité $d(x, a, Y_t^d)$ et à l'intensité d'évolution des traits $e(x, a, Y_t^e)$.

Il est possible de décrire la dynamique de ces facteurs aléatoires continus par un système d'équations différentielles stochastiques :

$$\begin{aligned} Y_t^b &= Y_0^b + \int_0^t \mu^b(s, Y_s^b) ds + \int_0^t \sigma^b(s, Y_s^b) dB_s^b, \\ Y_t^d &= Y_0^d + \int_0^t \mu^d(s, Y_s^d) ds + \int_0^t \sigma^d(s, Y_s^d) dB_s^d, \\ Y_t^e &= Y_0^e + \int_0^t \mu^e(s, Y_s^e) ds + \int_0^t \sigma^e(s, Y_s^e) dB_s^e. \end{aligned} \tag{6.4}$$

avec B_s^b , B_s^d et B_s^e des facteurs browniens vectoriels indépendants à valeurs respectivement dans \mathbb{R}^{N_b} , \mathbb{R}^{N_d} et \mathbb{R}^{N_e} . En prenant l'exemple du facteur de natalité Y_t^b , la tendance d'évolution μ^b est un vecteur à valeurs dans \mathbb{R}^{N_b} et la volatilité σ^b est une matrice à valeurs dans $\mathcal{M}_{N_b}(\mathbb{R})$.

En pratique, un schéma de discrétisation d'Euler leur est associé afin de simuler le processus d'évolution. Les données étant disponibles chaque année, nous considérons généralement la dynamique discrète suivante

$$\begin{aligned} Y_k^b &= Y_{k-1}^b + \mu^b(k-1, Y_{k-1}^b) + \sigma^b(k-1, Y_{k-1}^b)\epsilon^b, \\ Y_k^d &= Y_{k-1}^d + \mu^d(k-1, Y_{k-1}^d) + \sigma^d(k-1, Y_{k-1}^d)\epsilon^d, \\ Y_k^e &= Y_{k-1}^e + \mu^e(k-1, Y_{k-1}^e) + \sigma^e(k-1, Y_{k-1}^e)\epsilon^e, \end{aligned} \quad (6.5)$$

où ϵ^b , ϵ^d et ϵ^e sont des variables aléatoires gaussiennes centrées réduites. Les processus continus associés à ces facteurs peuvent s'écrire sous la forme $Y_t^b = Y_{[t]}^b$, $Y_t^d = Y_{[t]}^d$ et $Y_t^e = Y_{[t]}^e$ où $[x]$ correspond à la partie entière de x .

6.2.3 Hypothèses sur les intensités d'évolution

Il faut émettre des hypothèses sur les intensités d'évolution qui seront essentielles dans la définition du processus microscopique.

1. Tout d'abord, décrivons le caractère borné des intensités qui permet l'existence du processus d'évolution.

Supposons que les intensités b , d et e sont des processus continus et qu'il existe des constantes strictement positives \bar{b} , \bar{e} et \bar{d} telles que $\forall t \in \mathbb{R}_+$, $\forall (x, a) \in \tilde{\chi}$, nous avons presque sûrement :

$$|b(x, a, Y_t^b)| \leq \bar{b}, \quad |d(x, a, Y_t^d)| \leq \bar{d}, \quad |e(x, a, Y_t^e)| \leq \bar{e}.$$

2. De plus, supposons que les distributions des traits K^b et K^e lors des naissances et des évolutions des traits soient absolument continues par rapport à une probabilité de "référence" $P(dx)$ sur χ . Ainsi, les distributions $K^b(x, a, dx')$ et $K^e(x, a, dx')$ possèdent des densités bornées par des constantes positives $\bar{k}^b > 0$ et $\bar{k}^e > 0$. Ainsi, $\forall (x, a, x') \in \chi \times \mathbb{R}_+ \times \chi$, nous avons :

$$\begin{aligned} K^b(x, a, dx') &= k^b(x, a, x')P(dx') \quad \text{avec} \quad |k^b(x, a, x')| \leq \bar{k}^b, \\ K^e(x, a, dx') &= k^e(x, a, x')P(dx') \quad \text{avec} \quad |k^e(x, a, x')| \leq \bar{k}^e. \end{aligned}$$

Remarque 6.5. Lorsque $\chi \subset \mathbb{R}^d$, un choix naturel est la mesure de Lebesgue. Lorsque $\chi \subset \mathbb{N}^d$ et est borné, un choix naturel pourrait être la loi uniforme sur χ .

3. Enfin, il existe une hypothèse additionnelle de mortalité qui est assez classique dans l'évolution de population. Cette hypothèse consiste à supposer qu'il existe un taux de "mort naturelle" $\underline{d} : \tilde{\chi} \mapsto \mathbb{R}_+$ telle que presque sûrement, nous avons :

$$\begin{aligned} \forall (x, a) \in \tilde{\chi}, \quad \forall t \in \mathbb{R}_+, \quad d(x, a, Y_t^d) &\geq \underline{d}(x, a), \\ \forall (x, a) \in \tilde{\chi}, \quad \forall t \in \mathbb{R}_+, \quad \int_t^{+\infty} d(x, a + (u - t), Y_u^d) ds &= +\infty. \end{aligned}$$

Cette dernière hypothèse concernant la mortalité naturelle \underline{d} permet de démontrer que la probabilité de survie d'un individu tend vers 0 lorsque le temps tend vers l'infini ou autrement dit que la probabilité qu'un individu vive éternellement est nulle. Cette hypothèse n'est pas nécessaire dans l'étude théorique du processus mais elle permet d'avoir une notion biologique et réaliste du taux de mortalité.

6.2.4 Probabilité de survie

Intéressons nous à la durée de vie d'un individu ainsi qu'à la probabilité de survie d'un individu jusqu'à une date donnée. Cette notion est très importante et renseigne sur l'espérance de vie des individus d'un âge donné.

- Supposons dans un premier temps que le facteur $Y_u^d = y_u^d$ est déterministe et que les traits sont fixes dans le temps.

La durée de vie d'un individu né au temps $c \in \mathbb{R}_+$ et caractérisé par un trait $x \in \chi$. Il s'agit d'une variable aléatoire de densité D donnée par :

$$D : s \mapsto d(x, c + s, y_s^d) \exp \left(- \int_c^{c+s} d(x, c + u, y_u^d) du \right).$$

Cette expression est bien une densité de probabilité puisqu'il s'agit d'une application intégrable positive telle que :

$$\begin{aligned} \int_0^{+\infty} D(s) ds &= \int_0^\infty d(x, c + s, y_s^d) \exp \left(- \int_c^{c+s} d(x, c + u, y_u^d) du \right) ds \\ &= \left[- \exp \left(- \int_c^{c+s} d(x, c + u, y_u^d) du \right) \right]_0^{+\infty} = 1 \end{aligned}$$

étant donnée l'hypothèse sur la mortalité naturelle qui fournit

$$0 \leq \lim_{s \rightarrow +\infty} \exp \left(- \int_c^{c+s} d(x, c + u, y_u^d) du \right) \leq \lim_{s \rightarrow +\infty} \exp \left(- \int_c^{c+s} \underline{d}(x, c + u) du \right) = 0.$$

La probabilité de survie est la probabilité $S_{t_0}(x, a, t)$ qu'un individu de trait $x \in \chi$, d'âge $a \in \mathbb{R}_+$ au temps $t_0 \in \mathbb{R}_+$ soit toujours vivant à la date $t > t_0$ est donné par :

$$\begin{aligned} S_{t_0}(x, a, t) &= \int_t^\infty d(x, a + s - t_0, y_s^d) \exp \left(- \int_{t_0}^s d(x, a + u - t_0, y_u^d) du \right) ds \\ &= \exp \left(- \int_{t_0}^t d(x, a + u - t_0, y_u^d) du \right). \end{aligned}$$

L'hypothèse sur la mortalité naturelle permet de déduire que $\lim_{t \rightarrow +\infty} S_{t_0}(x, a, t) = 0$. Cela signifie que la probabilité qu'un individu vive éternellement est nulle.

- Dans notre cas, le facteurs Y_t^d est aléatoire. Une nouvelle expression naturelle de la probabilité de survie d'un individu est donnée par :

$$S_{t_0}(x, a, t) = \mathbb{E} \left[\exp \left(- \int_{t_0}^t d(x, a + u - t_0, Y_u^d) du \right) \middle| \mathcal{F}_{t_0} \right].$$

- De plus, nous pouvons rajouter le fait que les traits peuvent évoluer dans le temps avec un taux d'évolution de traits e . En considérant en date 0 un individu d'âge initial $a \in \mathbb{R}_+$ et de trait initial $x \in \chi$, nous définissons le processus $\tilde{X}_t^{x,a} = (X_t, A_t)^{x,a}$ où le processus $X_t^{x,a}$ correspond au trait de l'individu à la date t qui évolue à l'aide du taux d'évolution e et le processus $A_t^{x,a} = a + t$ correspond à l'âge de l'individu à l'instant t .

Ce processus \tilde{X} est une chaîne inhomogène de Markov en temps continu décrivant l'évolution des caractéristiques d'un individu dans le temps. Ainsi, si à l'instant t le processus $\tilde{X}_t^{a,x}$ est connu, l'âge évolue de manière déterministe dans le temps alors que le trait évolue au bout d'un temps exponentiel $e(X_t^{x,a}, A_t^{x,a}, Y_t^e)$. Les transitions de passage d'un trait à un autre dépendent du temps t et sont données par la densité de distribution de traits $K^e(X_t^{x,a}, A_t^{x,a}, dx')$.

Par conséquent, nous suggérons la nouvelle expression de la probabilité de survie qui prend en compte les éventuels changements de traits au cours de la vie de l'individu :

$$\begin{aligned} \tilde{X}_t^{x,a} &= (X_t^{x,a}, A_t^{x,a}), \\ S_{t_0}(X_{t_0}^{x,a}, A_{t_0}^{x,a}, t) &= \mathbb{E} \left[\exp \left(- \int_{t_0}^t d(X_u^{x,a}, A_u^{x,a}, Y_u^d) du \right) \middle| \mathcal{F}_{t_0} \right], \end{aligned}$$

où le processus markovien $\tilde{X}_t^{x,a}$ décrit l'évolution dans le temps de l'âge et des traits d'un individu d'âge a et de trait x en date 0.

6.2.5 Lien entre les notions de taux et d'intensité

La notion classique de taux en démographie est différente de celle d'intensité en dynamique de population. Prenons l'exemple de la mortalité afin de décrire la différence entre ces deux concepts.

La notion de taux de mortalité est généralement annuelle et correspond à la probabilité qu'un individu décède au cours d'une année. En effet, le taux de mortalité classique $q(x, a, Y_t^d)$ correspond à la probabilité d'un individu d'âge a et de trait x de décéder au cours de l'année. Ainsi, la probabilité de survie $S_t(x, a, t + h)$ d'un individu d'âge a et de trait x à la date t pendant une année est directement reliée au taux de mortalité :

$$S_t(x, a, t + 1) = 1 - q(x, a, Y_t^d).$$

L'intensité de mortalité $d(x, a, Y_t^d)$ correspond au paramètre du processus d'évolution qui correspond aussi au temps moyen entre deux événements. cette notion peut aussi être reliée à la fonction de survie. En effet, la probabilité de survie $S_t(x, a, t + 10)$ d'un individu d'âge a et de trait x à la date t jusqu'à la date $t + 1$ peut s'exprimer en fonction de l'intensité de mortalité $d(a, x, Y_t^d)$ sous la forme suivante

$$S_t(x, a, t + 1) = \mathbb{E} \left[\exp \left(- \int_t^{t+1} d(x, a - t + u, Y_u^d) du \right) \middle| \mathcal{F}_t \right].$$

En faisant l'approximation d'une intensité de mortalité d constante sur une année et en identifiant les deux expressions de la fonction de survie, nous obtenons :

$$d(x, a, Y_t^d) \sim \ln \left(\frac{1}{1 - q(x, a, Y_t^d)} \right).$$

Cette relation fournit le lien entre une intensité de mortalité et un taux de mortalité et plus généralement entre une intensité d'évolution et un taux d'évolution. En pratique, c'est souvent le taux d'évolution qui est modélisé plutôt que l'intensité d'évolution dans la mesure où le taux apparaît plus facile à analyser. Lorsque les taux sont petits, un développement de Taylor permet de remarquer que le taux de mortalité q et l'intensité de mortalité d ont des valeurs très proches.

6.3 Processus microscopique

En Section 6.1.3, nous avons introduit le concept de processus ponctuel discret d'évolution qui est essentiel dans une modélisation microscopique d'évolution de population. La prise en compte de l'âge demande de raffiner le modèle dans le sens proposé par N. Fournier et S. Méléard [18] et de V.C. Tran [39, 40]. Nous considérons le processus d'évolution comme un processus stochastique, à valeurs dans un espace de mesures ponctuelles, ayant des propriétés de régularité trajectorielle. Il nous faut tout d'abord rappeler quelques notions mathématiques.

6.3.1 L'espace des mesures ponctuelles

Pour deux espaces métriques E et F , notons $\mathcal{C}_b(E, F)$ l'ensemble des fonctions continues bornées de E dans F qui est muni de la norme uniforme. Lorsque $F = \mathbb{R}$, nous utiliserons la notation $\mathcal{C}_b(E)$. L'ensemble $\mathbb{D}(E, F)$ des fonctions càd-làg de E dans F est muni de la distance de Skorokhod.

Notons $\mathcal{M}_P(E)$ l'ensemble des mesures ponctuelles sur E que nous munissons de la topologie de la convergence faible des mesures. Une mesure ponctuelle ν sur E peut s'écrire sous la forme :

$$\nu = \sum_n \delta_{e_n},$$

où δ_a désigne la mesure de Dirac au point a et où e_n est une suite d'éléments de H .

Pour $m \in \mathcal{M}_P(E)$ et $f \in \mathcal{B}_b(E)$ l'ensemble des fonctions boréliennes bornées, nous notons $\langle m, f \rangle = \int_E f dm$.

L'espace des mesures finies sur E , noté $\mathcal{M}_F(E)$, peut être muni de la convergence étroite ou de la convergence vague. Nous le munirons généralement de la convergence étroite mais lorsqu'il y aura une ambiguïté, nous préciserons de quel type de convergence il s'agit, $(\mathcal{M}_F(E), \epsilon)$ ou $(\mathcal{M}_F(E), v)$.

6.3.2 Définition du processus d'évolution

Les individus que nous considérons seront caractérisés par leurs traits correspondant à un ensemble de caractéristiques appartenant à un espace métrique χ muni d'une mesure de référence P . Nous distinguons, le paramètre de l'âge, à valeurs dans \mathbb{R}_+ et notons $\tilde{\chi} = \chi \times \mathbb{R}_+$. Un individu sera donc caractérisé par une mesure de Dirac en $(x, a) \in \tilde{\chi}$. À tout instant t , la population peut être décrite par la mesure de $\mathcal{M}_P(\tilde{\chi})$ donnée par :

$$Z_t = \sum_{i=1}^{N_t} \delta_{(x_i(t), a_i(t))}, \quad \text{où} \quad N_t = \langle Z_t, 1 \rangle = \int_{\tilde{\chi}} Z_t(dx, da) \quad (6.6)$$

est le nombre d'individus à l'instant t .

Introduisons le processus d'évolution d'une population structurée par âge et par traits qui généralise la modélisation proposée par l'expression (6.4) dans la section 6.1.3 aux populations structurées par traits. Les notations de la section 6.1.3 seront reprises sans rappel. Les propriétés des processus d'intensités sont détaillées dans la section 6.2.3.

Définition 6.3. Sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, définissons un ensemble d'éléments aléatoires indépendants de la manière suivante :

- La population initiale $Z_0 \in \mathcal{M}_P(\tilde{\chi})$ de taille finie est une mesure ponctuelle telle que $\mathbb{E}(\langle Z_0, 1 \rangle) < \infty$.
- $Q(ds, di, d\theta, dx') = \sum_{n \geq 1} \delta_{(S_n, I_n, \Theta_n, X'_n)}(ds, di, d\theta, dx')$ est une mesure ponctuelle de Poisson sur l'espace $\mathbb{R}_+ \times \Xi = \mathbb{R}_+ \times \mathbb{N}^* \times \mathbb{R}_+ \times \chi$ dont le compensateur est $ds n(di) d\theta P(dx')$.
- Les facteurs exogènes Y_t^b , Y_t^d et Y_t^e sont des facteurs communs à l'ensemble de la population (à l'échelle de la population et non pas à l'échelle de l'individu), qui sont indépendants de Q , de Z_0 et qui sont indépendants entre eux.

Notons $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ la filtration canonique engendrée par Z_0 , Q et par les facteurs (Y^b, Y^d, Y^e) régularisée. Nous définissons un processus de dynamique de population Z à valeurs dans les mesures ponctuelles sur $\tilde{\chi}$ par son évolution :

$$\begin{aligned}
 Z_t = & \sum_{i=0}^{N_0} \delta_{(X_i(Z_0), A_i(Z_0)+t)} + \int_0^t \int_{\Xi} 1_{i \leq N_{s-}} \left[\delta_{(x', t-s)} 1_{0 \leq \theta < m_1(s, Z_{s-}, i, x')} \right. \\
 & - \delta_{(X_i(Z_{s-}), A_i(Z_{s-})+t-s)} 1_{m_1(s, Z_{s-}, i, x') \leq \theta < m_2(s, Z_{s-}, i, x')} \\
 & \left. + (\delta_{(x', A_i(Z_{s-})+t-s)} - \delta_{(X_i(Z_{s-}), A_i(Z_{s-})+t-s)}) 1_{m_2(s, Z_{s-}, i, x') \leq \theta < m_3(s, Z_{s-}, i, x')} \right] Q(ds, di, d\theta, dx'),
 \end{aligned} \tag{6.7}$$

où $X_i(Z_t)$ et $A_i(Z_t)$ correspondent respectivement aux traits et à l'âge de l'individu i de la population à la date t et où :

$$\begin{aligned}
 m_1(s, Z_{s-}, i, x') &= b(X_i(Z_{s-}), A_i(Z_{s-}), Y_s^b) k^b(X_i(Z_{s-}), A_i(Z_{s-}), x'). \\
 m_2(s, Z_{s-}, i, x') &= m_1(s, Z_{s-}, i, x') + d(X_i(Z_{s-}), A_i(Z_{s-}), Y_s^d) k^d(X_i(Z_{s-}), A_i(Z_{s-}), x'). \\
 m_3(s, Z_{s-}, i, x') &= m_2(s, Z_{s-}, i, x') + e(X_i(Z_{s-}), A_i(Z_{s-}), Y_s^e) k^e(X_i(Z_{s-}), A_i(Z_{s-}), x').
 \end{aligned}$$

Par conséquent, le processus Z est un processus càd-làg à valeurs dans l'espace des mesures ponctuelles, qui entre deux sauts n'évolue que par l'influence de l'âge. Les événements affectant la population sont les naissances et les décès ainsi que l'évolution des traits des individus. Les m_i représentent les intensités individuelles d'évolution (issues de processus de Poisson) et permettent de caractériser les occurrences des événements d'évolution. Entre deux événements, les individus vieillissent continûment avec le temps. Le rôle de la mesure ponctuelle canonique $Q(ds, di, d\theta, dx')$ est similaire à celui du mouvement brownien dans une équation différentielle stochastique.

Le terme $\sum_{i=0}^{N_0} \delta_{(X_i(Z_0), A_i(Z_0)+t)}$ correspond à la structure de la population à la date t s'il n'y avait eu que le vieillissement. Les autres termes permettent de tenir compte des différents événements qui modifient la structure de la population. L'ensemble des masses de Dirac qui décrivent la population à la date t est constitué des mesures de Dirac représentant les individus présents initialement auxquelles s'ajoutent les mesures de Dirac représentant les individus nés en un temps $s \in [0, t]$ et auxquelles sont soustraites les mesures de Dirac représentant les individus morts en un temps $s \in [0, t]$. Le dernier terme correspond à d'éventuels changements de traits dans le temps, ce qui correspond à remplacer la masse de Dirac par une autre ayant le même âge mais un trait différent.

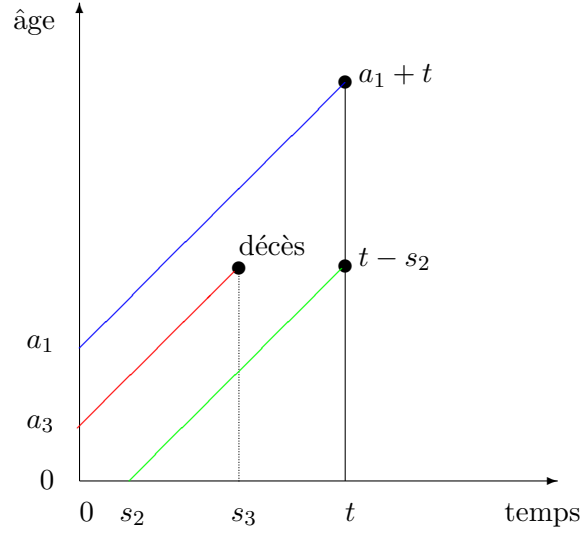


FIGURE 6.1 – Représentation de la structure des âges en fonction du temps en tenant compte de différents événements. La ligne bleue correspond à un individu d'âge initial a_1 et âgé de $a_1 + t$ en date t . La ligne verte correspond à un individu né en date s_2 qui est âgé de $t - s_2$ en date t . La ligne rouge correspond à l'âge d'un individu d'âge initial a_3 qui disparaît en date s_3 .

Le processus d'évolution est décrit sous la forme d'une mesure, mais nous pouvons aussi déterminer naturellement le processus d'évolution par $\langle Z_t, f_t \rangle$ sur un ensemble de fonctions tests $f : (x, a, s) \mapsto f_s(x, a) \in \mathcal{B}_b(\tilde{\mathcal{X}} \times \mathbb{R}_+)$:

$$\begin{aligned} \langle Z_u, f_u \rangle &= \langle Z_0, f_0 \rangle + \int_0^u \int_{\Xi} \left(f_s(x', 0) 1_{0 \leq \theta < m_1(s, Z_{s-}, i, x')} \right. \\ &\quad - f_s(X_i(Z_{s-}), A_i(Z_{s-})) 1_{m_1(s, Z_{s-}, i, x') \leq \theta < m_2(s, Z_{s-}, i, x')} \\ &\quad \left. + \left[f_s(x', A_i(Z_{s-})) - f_s(X_i(Z_{s-}), A_i(Z_{s-})) \right] 1_{m_2(s, Z_{s-}, i, x') \leq \theta < m_3(s, Z_{s-}, i, x')} \right) \\ &\quad Q(ds, di, d\theta, dx') \end{aligned}$$

Cette forme ne fait pas intervenir les masses de Dirac représentant les individus. L'interprétation ne se fait plus sous forme de mesures mais plutôt sous la forme de comptage qui est la forme standard des processus de naissance et mort.

Les hypothèses consistant à supposer que les intensités individuelles sont bornées sont assez naturelles, en particulier dans le cadre de l'étude d'une population humaine. L'introduction en pratique des facteurs dans les intensités d'évolution ne changera pas le fait que les fonctions b , d et e seront bornées et même comprises entre 0 et 1. En effet, la modélisation des taux d'évolution utilisant la fonction logit dont l'inverse permet de récupérer des taux

démographiques bornés entre 0 et 1 (voir Chapitre 5). L'hypothèse sur le caractère borné de ces intensités est essentielle puisqu'elle permet d'assurer l'existence et l'unicité forte du processus microscopique (voir Proposition 6.1) ainsi qu'un procédé de simulation efficace.

6.3.3 Simulation du processus d'évolution

Nous reprenons dans cette partie la méthode de simulation introduite par S. Méléard et N. Fournier [18] de processus microscopiques et adaptée par V.C Tran [39] pour traiter l'évolution d'une population structurée par âge. Nous adaptons ce procédé de simulation à ce cadre en ajoutant l'évolution des traits des individus et en supposant que les intensités d'évolution dépendent des facteurs Y^b , Y^d et Y^e . Les intensités de natalité, de mortalité et d'évolution des traits sont supposées majorées respectivement par \bar{b} , \bar{d} et \bar{e} .

Les facteurs Y^d , Y^b et Y^e évoluent suivant la dynamique (6.4) et leur simulation se fait par un schéma d'Euler avec un pas de discrétisation h . Si nous privilégions une évolution discrète par année avec l'expression (6.5), cela revient aussi à fixer le taux de discrétisation à $h = 1$.

Considérons la date d'occurrence du k^e événement $T_k \in [nh, (n+1)h]$ où n correspond au nombre de dates du schéma de discrétisation de Euler de pas h depuis la date initiale. Dans une description d'évolution annuelle avec $h = 1$, n correspondra au 1er janvier de l'année au cours de laquelle se produit le k^e événement et $n+1$ correspond au premier janvier de l'année suivante.

Ainsi, le taux de saut global du processus est majoré par $(\bar{b} + \bar{d} + \bar{e})N_{T_k}$ et dépend de la taille de la population puisqu'on sélectionne un individu dans la population de taille N_{T_k} .

Remarque 6.6. *Si nous avons modélisé une mortalité par compétition [39], la majoration serait en $\bar{b}N_{T_k} + \bar{d}f(N_{T_k}) + \bar{e}N_{T_k}$ ce qui traduit une interaction des individus.*

On détermine le temps T_{k+1} du prochain événement par une méthode d'acceptation/rejet. Considérons la suite (τ_{k+l}) de processus ponctuel de Poisson d'intensité $(\bar{b} + \bar{d} + \bar{e})N_{T_k}$ et mettons en place la procédure d'acceptation rejet.

1. On note $\tau_k = T_k \in [nh, (n+1)h]$, $N_{\tau_k} = N_{T_k}$, $Y^d = Y_{nh}^d$, $Y^b = Y_{nh}^b$, $Y^e = Y_{nh}^e$, $l := 0$.
2. On détermine le prochain instant auquel un événement est susceptible de se produire. On tire une variable aléatoire $\epsilon_{k,l} \sim \mathcal{E}((\bar{b} + \bar{d} + \bar{e})N_{T_k})$ ce qui fournit un instant éventuel de saut donné par $\tau_{k+l+1} = \tau_{k+l} + \epsilon_{k,l}$. Si $\tau_{k+l+1} > (n+1)h$, on simule $Y^b = Y_{(n+1)h}^b$, $Y^d = Y_{(n+1)h}^d$ et $Y^e = Y_{(n+1)h}^e$ à partir de Y_{nh}^b , Y_{nh}^d et Y_{nh}^e .

Remarque 6.7. *Le pas de temps consiste à tirer une variable exponentielle de paramètre $(\bar{b} + \bar{d} + \bar{e})N_{T_k}$. En fait, on peut aussi interpréter cela comme le fait d'asso-*

cier à chaque individu une cloche sonnante suivant la loi exponentielle de paramètre $(\bar{b} + \bar{d} + \bar{e})$. Le fait de regarder la première son de cloche revient à prendre la borne inférieure de N_{T_k} lois exponentielles indépendantes identiquement distribuées associées à l'ensemble cloche. Cela correspond à considérer une loi exponentielle dont le paramètre est la somme des paramètres. Ainsi, on retrouve bien le pas de temps $\epsilon_{k,l}$ proposé.

3. Pendant la période $[\tau_{k+l}, \tau_{k+l+1}]$, seul le phénomène de vieillissement apparaît et un individu de la population est caractérisé au temps τ_{k+l+1} par son trait $X_i(Z_{\tau_{k+l}})$ et son âge $A_i(Z_{\tau_{k+l}}) + \tau_{k+l+1} - \tau_{k+l}$.
4. On sélectionne au hasard un individu dans la population Z_{T_k} ce qui revient à tirer une variable aléatoire $I_{k,l}$ uniformément distribuée sur $[1, N_{T_k}]$. On calcule les valeurs des différents \tilde{m}_i dans $[0, 1]$ par :

$$\begin{aligned}\tilde{m}_1(\tau_{k+l+1}, Z_{\tau_{(k+l+1)}^-}, I_{k,l}) &= \frac{b(X_{I_{k,l}}(Z_{\tau_{(k+l+1)}^-}), A_{I_{k,l}}(Z_{\tau_{(k+l+1)}^-}), Y^b)}{\bar{b} + \bar{d} + \bar{e}}. \\ \tilde{m}_2(\tau_{k+l+1}, Z_{\tau_{(k+l+1)}^-}, I_{k,l}) &= \tilde{m}_1(\tau_{k+l+1}, Z_{\tau_{(k+l+1)}^-}, I_{k,l}) + \frac{d(X_{I_{k,l}}(Z_{\tau_{(k+l+1)}^-}), A_{I_{k,l}}(Z_{\tau_{(k+l+1)}^-}), Y^d)}{\bar{b} + \bar{d} + \bar{e}}. \\ \tilde{m}_3(\tau_{k+l+1}, Z_{\tau_{(k+l+1)}^-}, I_{k,l}) &= \tilde{m}_2(\tau_{k+l+1}, Z_{\tau_{(k+l+1)}^-}, I_{k,l}) + \frac{e(X_{I_{k,l}}(Z_{\tau_{(k+l+1)}^-}), A_{I_{k,l}}(Z_{\tau_{(k+l+1)}^-}), Y^e)}{\bar{b} + \bar{d} + \bar{e}}.\end{aligned}$$

5. On détermine le type d'événement qui va affecter l'individu sélectionné. Pour cela, on tire une variable $\theta_{k,l} \sim U_{[0,1]}$
 - si $0 < \theta_{k,l} < \tilde{m}_1$, l'individu $I_{k,l}$ donne naissance à un individu d'âge 0 et de trait x' tiré suivant la loi $K^b(X_{I_{k,l}}(Z_{\tau_{(k+l+1)}^-}), A_{I_{k,l}}(Z_{\tau_{(k+l+1)}^-}), dx')$, on définit $T_{k+1} = \tau_{k+l+1}$ et on revient à l'étape 1.
 - si $\tilde{m}_1 < \theta_{k,l} < \tilde{m}_2$, l'individu $I_{k,l}$ meurt, on définit $T_{k+1} = \tau_{k+l+1}$ et on revient à l'étape 1.
 - si $\tilde{m}_2 < \theta_{k,l} < \tilde{m}_3$, l'individu $I_{k,l}$ change de trait et passe du trait $X_{I_{k,l}}(Z_{\tau_{(k+l+1)}^-})$ à un trait x' tiré suivant la loi $K^e(X_{I_{k,l}}(Z_{\tau_{(k+l+1)}^-}), A_{I_{k,l}}(Z_{\tau_{(k+l+1)}^-}), dx')$, on définit $T_{k+1} = \tau_{k+l+1}$ et on revient à l'étape 1.
 - si $\theta_{k,l} > \tilde{m}_3$, aucun événement ne se produit (ou plutôt un événement auxiliaire qui n'affecte pas l'évolution de la population) et on pose $N_{\tau_{k+l+1}} = N_{T_k}$, $l \rightarrow l + 1$ et on revient à l'étape 2.

On observe que les pas de temps pour la simulation du processus suivent des lois exponentielles dont l'intensité à chaque pas dépend de la taille de la population. On perd

la propriété de temps de Poisson puisqu'ils dépendent de l'évolution de la structure de la population. Numériquement, la simulation du processus se fait par une discrétisation avec un pas qui est aléatoire et adaptatif. Cette méthode de simulation permet de sélectionner de manière optimale l'ensemble des individus de la population. On propose dans l'annexe A un exemple simple expliquant la renormalisation (ou "Thinning") qui permet d'avoir un procédé de simulation d'une mesure de Poisson sur un espace de mesure infinie.

Remarque 6.8. *On aurait pu conserver la structure poissonnienne en limitant la taille de la population à un seuil \bar{N} : on aurait tiré une variable exponentielle $\epsilon^{k,l}$ de paramètre $\bar{N}(\bar{b} + \bar{d} + \bar{e})$, on aurait tiré une variable $I_{k,l}$ uniformément distribuée sur $[1, \bar{N}]$ et on aurait rejeté les valeurs supérieures à N_{T_k} . Cependant, cette simulation bien que poissonnienne est moins efficace puisqu'on sélectionne un trop grand nombre d'événements dont la plupart sont rejetés.*

En théorie, le procédé de simulation est indépendant des valeurs des bornes \bar{b} , \bar{d} et \bar{e} . En pratique, le modèle est d'autant plus efficace lorsque ces bornes ont des valeurs proches des vraies bornes supérieures des intensités d'évolution afin de limiter le nombre de rejets dans la procédure. Ainsi, l'idée est donc de simuler une trajectoire des facteurs d'évolution Y^b , Y^d et Y^e et de considérer les bornes optimales des intensités d'évolution associées à cette trajectoire. Une fois les bornes \bar{b} , \bar{d} et \bar{e} déterminées, on peut commencer à simuler le processus d'évolution de la population sachant cette trajectoire.

6.3.4 Migration et immigration

Nous proposons trois méthodes différentes afin de prendre en compte les migrations et l'immigration dans le modèle :

- Intensités individuelles d'immigration et de migration
- Intensité individuelle de migration et intensité globale d'immigration
- Intensité globale d'immigration en discret (par simulation)

Migration et immigration individuelles

Les migrations et l'immigration dans la modélisation peuvent être pris en compte de sorte à avoir des flux continus d'individus selon une intensité d'immigration individuelle.

Au même titre que les autres intensités d'évolution, nous ajoutons une intensité $\rho^-(x, a, Y_t^-)$ correspondant à une intensité individuelle de migration de la population. et une intensité $\rho^+(x, a, Y_t^+)$ correspond à une intensité d'immigration individuelle. L'intensité ρ^+ correspond à une intensité d'immigration d'un individu d'âge a et de trait x qui a le même âge que l'individu sélectionné et dont le trait est tiré selon la distribution $k^I(x, a, x')$.

Pour information, le processus d'évolution modifié se réécrit :

$$\begin{aligned}
 Z_t^I &= \sum_{i=0}^{N_0} \delta_{(X_i(Z_0), A_i(Z_0)+t)} \\
 &+ \int_0^t \int_{\Xi} 1_{i \leq N_{s-}} \left[\delta_{(x', t-s)} 1_{0 \leq \theta < m_1(s, Z_{s-}, i, x')} \right. \\
 &\quad - \delta_{(X_i(Z_{s-}), A_i(Z_{s-})+t-s)} 1_{m_1(s, Z_{s-}, i, x') \leq \theta < m_2(s, Z_{s-}, i, x')} \\
 &\quad + (\delta_{(x', A_i(Z_{s-})+t-s)} - \delta_{(X_i(Z_{s-}), A_i(Z_{s-})+t-s)}) 1_{m_2(s, Z_{s-}, i, x') \leq \theta < m_3(s, Z_{s-}, i, x')} \\
 &\quad - \delta_{(X_i(Z_{s-}), A_i(Z_{s-})+t-s)} 1_{m_3(s, Z_{s-}, i, x') \leq \theta < m_4(s, Z_{s-}, i, x')} \\
 &\quad \left. + \delta_{(x', A_i(Z_{s-})+t-s)} 1_{m_4(s, Z_{s-}, i, x') \leq \theta < m_5(s, Z_{s-}, i, x')} \right] Q(ds, di, d\theta, dx'), \quad (6.8)
 \end{aligned}$$

où les coefficients m_i sont donnés par

$$\begin{aligned}
 m_1(s, Z_{s-}, i, x') &= b(X_i(Z_{s-}), A_i(Z_{s-}), Y_s^b) k^b(X_i(Z_{s-}), A_i(Z_{s-}), x'). \\
 m_2(s, Z_{s-}, i, x') &= m_1(s, Z_{s-}, i, x') + d(X_i(Z_{s-}), A_i(Z_{s-}), Y_s^d) k^b(X_i(Z_{s-}), A_i(Z_{s-}), x'). \\
 m_3(s, Z_{s-}, i, x') &= m_2(s, Z_{s-}, i, x') + e(X_i(Z_{s-}), A_i(Z_{s-}), Y_s^e) k^e(X_i(-Z_{s-}), A_i(Z_{s-}), x'). \\
 m_4(s, Z_{s-}, i, x') &= m_3(s, Z_{s-}, i, x') + \rho^-(X_i(Z_{s-}), A_i(Z_{s-}), Y_s^-) k^I(X_i(-Z_{s-}), A_i(Z_{s-}), x'). \\
 m_5(s, Z_{s-}, i, x') &= m_4(s, Z_{s-}, i, x') + \rho^+(X_i(Z_{s-}), A_i(Z_{s-}), Y_s^+) k^I(X_i(-Z_{s-}), A_i(Z_{s-}), x').
 \end{aligned}$$

Remarque 6.9. *L'avantage majeur de ce formalisme pour le processus d'évolution est qu'il est aisé d'ajouter des événements sans pour autant modifier la structure générale du processus microscopique.*

Migration individuelle et immigration globale continue

Nous pouvons prendre en compte l'immigration dans la modélisation à une échelle macroscopique et non plus microscopique. Cela permet d'avoir une arrivée continue d'immigrants selon une intensité d'immigration globale dépendant uniquement de la taille de la population. Nous conservons une intensité de migration individuelle $\rho^-(x, a, Y_t^-)$ mais nous passons à une intensité d'immigration globale $\rho^+(Z_s, Y_s)$ ne dépendant que du temps et de la taille de la population. L'âge et les traits de l'immigré sont tirés selon une loi définie au préalable de densité $k^I(x, a)$.

Ces deux premières modélisations contraignent l'immigration à être modélisée alors qu'en pratique nous pourrions souhaiter adapter une politique d'immigration avec le temps en fonction de l'évolution de la population. Nous exhibons une troisième méthode permettant de mettre en place des stratégies d'immigration.

Migration individuelle et immigration discrète

Dans cette dernière méthode, l'intensité d'immigration peut correspondre à une stratégie adaptative dépendant de certains critères (comme nous le verrons dans le chapitre 7 de la thèse). L'intensité d'immigration est déterminée à certaines dates fixées (chaque année par exemple) en fonction de l'évolution de la structure de la population. Par conséquent, l'immigration se fait de manière discrète chaque année au cours de la simulation du processus et dépend du scénario d'évolution.

Cette méthode permet de tenir compte de l'immigration en arrêtant la simulation le processus (chaque année par exemple) et en mettant à jour la population en tenant compte des migrations et des immigrations à l'aide d'un critère d'immigration. Ce critère d'immigration pourra être lié à une stratégie d'immigration qui pourra dépendre de la taille de la population ou d'un critère purement arbitraire. Après avoir effectué les différents flux migratoires, on relance le processus de diffusion avec la population modifiée par l'immigration comme nouvelle condition initiale.

Conclusions sur l'immigration

Il est possible de définir un taux d'immigration par de nombreuses méthodes qui ont chacune leurs avantages et leurs inconvénients. L'immigration peut être fixée à l'avance par un taux individuel ou un taux global. Il est aussi possible de considérer des stratégies d'immigration visant à satisfaire un objectif donné. Dans la suite, par souci de simplicité, nous omettons ces termes d'immigration afin de ne pas alourdir les notations. Ainsi, l'immigration peut facilement être prise en compte dans la modélisation ou dans la simulation du processus.

6.3.5 Extensions possibles

Il existe de nombreuses extensions avec la prise en compte d'autres phénomènes concernant la structure et la dynamique de la population.

1. Nous avons considéré l'âge physique de l'individu qui est directement proportionnel au temps ce qui signifie que l'âge au temps t d'un individu né au temps c est donné par $t - c$. Il est possible de modéliser un âge biologique a qui évolue avec une vitesse $v(x, a)$ dans le temps [39, 40]. Cela signifie que l'âge d'un individu de trait x et d'âge a vieillit à une vitesse $v(x, a)$. À la date t , l'âge biologique d'un individu de trait x' et d'âge a à la date $t - 1$ est donnée par $a + v(x, a)$. La notion d'âge biologique pourrait être une manière astucieuse de tenir compte des maladies donnant l'impression qu'un individu vieillit plus rapidement qu'un individu en bonne santé.
2. Nous n'avons pas modélisé d'éventuelle interaction entre les individus. Il est possible de tenir compte d'une compétition entre les individus qui affecte les intensités in-

dividuelles de mortalité et de natalité [39, 40, 30]. Cette interaction complexifie la structure de la population et pourrait fournir des résultats très intéressants.

6.4 Propriétés mathématiques du processus

Dans cette partie, nous décrivons les différentes propriétés mathématiques relatives à ce processus d'évolution. Tout d'abord, nous prouvons l'existence et l'unicité du processus d'évolution. Nous démontrons des propriétés de contrôle du processus avec des propriétés d'existence de moments. Ensuite, nous caractérisons le processus par son générateur infinitésimal permettant l'analyse de la dynamique du processus au travers d'un problème martingale vérifié par ce dernier. Enfin, nous exhibons des propriétés de régularité avec l'existence d'une densité par âge et par traits qui permet de structurer la projection moyenne de la population.

6.4.1 Existence et unicité forte du processus

Établissons l'existence et l'unicité du processus d'évolution. Ce processus d'évolution est caractérisé par les événements d'évolution et il faut donc assurer l'existence et l'unicité des instant T_k d'évolution à toute date t .

Bien que l'évolution des traits ne change pas la taille effective de la population, elle en modifie la structure. Ainsi, le taux d'occurrence d'événements au temps $s \in \mathbb{R}_+$ est majoré par $(\bar{b} + \bar{d} + \bar{e})N_{s-}$. Il est donc possible de définir presque sûrement la suite $(T_k)_{k \in \mathbb{N}^*}$ des temps de événements successifs de Z . Comme il s'agit d'une suite croissante, $T_\infty = \lim_{k \rightarrow +\infty} T_k$ est aussi définie presque sûrement. Pour démontrer l'existence de la solution, il faut démontrer que $T_\infty = +\infty$ \mathbb{P} -presque sûrement.

Proposition 6.1. *Sous les hypothèses énoncées précédemment, l'équation (6.8) admet une solution $(Z_t)_{t \in \mathbb{R}_+} \in \mathbb{D}(\mathbb{R}_+, \mathcal{M}_P(\tilde{\chi}))$ telle que pour tout $T > 0$, N_t défini en (6.6) satisfait :*

$$\mathbb{E}\left(\sup_{t \in [0, T]} N_t\right) < \mathbb{E}(N_0) e^{\bar{b}T} < +\infty. \quad (6.9)$$

Démonstration 6.1. Soit $N \in \mathbb{N}^*$ fixé. On introduit le temps d'arrêt suivant :

$$\tau_N = \inf\{t \geq 0, N_t \geq N\}.$$

Pour tout $t \in \mathbb{R}_+$, on peut écrire :

$$\begin{aligned} \langle Z_{t \wedge \tau_N}, 1 \rangle &= N_0 + \int_0^{t \wedge \tau_N} \int_{\tilde{\chi}} 1_{\{i \leq N_{s-}\}} (1_{\{0 \leq \theta < m_1(s, Z_{s-}, i, x')\}} \\ &\quad - 1_{\{m_1(s, Z_{s-}, i, x') \leq \theta < m_2(s, Z_{s-}, i, x')\}}) Q(ds, di, d\theta, dx') \\ &\leq N_0 + \int_0^{t \wedge \tau_N} \int_{\tilde{\chi}} 1_{\{i \leq N_{s-}\}} 1_{\{0 \leq \theta \leq m_1(s, Z_{s-}, i, x')\}} Q(ds, di, d\theta, dx'). \end{aligned}$$

En prenant la borne supérieure et l'espérance, il est licite d'appliquer le théorème de Fubini (les intégrants sont bornés en considérant les hypothèses de la section 6.2.3 et le choix de τ_N). On déduit que

$$\mathbb{E}\left(\sup_{s \in [0, t \wedge \tau_N]} N_s\right) \leq \mathbb{E}(N_0) + \bar{b} \int_0^t \mathbb{E}\left(\sup_{u \in [0, s \wedge \tau_N]} N_u\right) ds \leq \mathbb{E}(N_0) e^{\bar{b}t}, \quad (6.10)$$

en utilisant le lemme de Gronwall. La majoration se faisant par un terme indépendant de N , on démontre par l'absurde que

$$\lim_{N \rightarrow +\infty} \tau_N = +\infty \quad \mathbb{P} \text{ p.s.}$$

En effet, s'il existait $M < +\infty$ et un ensemble $A_M \subset \Omega$ de probabilité strictement positive tels que $\forall \omega \in A_M, \lim_{N \rightarrow +\infty} \tau_N(\omega) < M$, alors,

$$\forall T > M, \quad \mathbb{E}\left(\sup_{t \in [0, T \wedge \tau_N]} N_t\right) \geq \mathbb{P}(A_M) \times N.$$

Ceci est en contradiction avec la majoration par un terme indépendant de N dans l'expression (6.10). Ainsi, $\lim_{N \rightarrow +\infty} \tau_N = +\infty \quad \mathbb{P}-\text{p.s.}$

Par conséquent, on peut utiliser le lemme de Fatou dans l'expression (6.10) et on obtient :

$$\mathbb{E}\left(\liminf_{N \rightarrow +\infty} \sup_{s \in [0, t \wedge \tau_N]} N_s\right) \leq \liminf_{N \rightarrow +\infty} \mathbb{E}\left(\sup_{s \in [0, t \wedge \tau_N]} N_s\right) \leq \mathbb{E}(N_0) e^{\bar{b}t} < +\infty,$$

ce qui prouve (6.9). À présent, démontrons par l'absurde que $T_\infty = +\infty \quad \mathbb{P}-\text{p.s.}$ ce qui assurera l'existence du processus. Supposons qu'il existe $\tilde{M} < +\infty$ et un ensemble $A_{\tilde{M}} \subset \Omega$ de probabilité strictement positive tels que $\forall \omega \in A_{\tilde{M}}, T_\infty(\omega) < \tilde{M}$.

Démontrons tout d'abord par l'absurde que

$$\forall \omega \in A_{\tilde{M}}, \quad \lim_{k \rightarrow +\infty} N_{T_k}(\omega) = +\infty. \quad (6.11)$$

Si la propriété (6.11) n'était pas vérifiée, il existerait $N' > 0$ et un ensemble $B \subset A_{\tilde{M}}$ de probabilité strictement positive tels que :

$$\forall \omega \in B, \forall k \in \mathbb{N}, N_{T_k}(\omega) < N'.$$

Ainsi, pour tout $\omega \in B$, on peut construire les instants $(T_k(\omega))_{k \in \mathbb{N}}$ comme une sous suite des points d'un processus ponctuelle de Poisson d'intensité $(\bar{b} + \bar{d} + \bar{e})N'$ dont la seule valeur d'adhérence est $+\infty \quad \mathbb{P}-$ presque sûrement ce qui est absurde car on avait supposé que $\forall \omega \in B, T_\infty(\omega) < \tilde{M}$.

La propriété (6.11) étant démontrée, on en déduit que $\forall N > 0, \forall \omega \in A_{\tilde{M}}, \tau_N(\omega) \leq \tilde{M}$ ce qui contredit le fait que $\lim_{N \rightarrow +\infty} \tau_N = +\infty \quad \mathbb{P}-\text{p.s.}$

Ceci achève la démonstration de l'existence d'un tel processus. \square

Proposition 6.2. *On a unicité forte pour l'équation différentielle stochastique (6.8)*

Démonstration 6.2. Soient $Z_0 \in \mathcal{M}_P(\tilde{\chi})$ une condition initiale et $Q(ds, di, d\theta, dx')$ une mesure ponctuelle de Poisson satisfaisant aux hypothèses de la définition 1. Démontrons par récurrence que les couples $(T_k, Z_{T_k})_{k \in \mathbb{N}}$ sont déterminés de façon unique par (Z_0, Q) . Pour $k = 0$, la propriété est évidente car Z_0 est déterminé de façon unique par le couple (Z_0, Q) . Supposons que pour $k \in \mathbb{N}$, (T_k, Z_{T_k}) soit déterminé de façon unique par (Z_0, Q) et démontrons que $(T_{k+1}, Z_{T_{k+1}})$ l'est aussi. On définit :

$$\forall t \geq T_k, \quad \tilde{Z}_t^k(dx, da) = \sum_{i=1}^{N_{T_k}} \delta_{(X_i(Z_{T_k}), A_i(Z_{T_k}) + t - T_k)}(dx, da)$$

qui est déterminé de façon unique par (T_k, Z_{T_k}) donc par (Z_0, Q) . Ainsi,

$$T_{k+1} = \inf \left\{ t > T_k, \int_{T_k}^t \int_{\tilde{\chi}} 1_{\{i \leq N_{T_k}, 0 \leq m_3(s, \tilde{Z}_{s-}^k, i, x')\}} Q(ds, di, d\theta, dx') > 0 \right\}$$

est déterminé de façon unique par (Z_0, Q) . Soit l'atome $(T_{k+1}, I_{k+1}, \theta_{k+1}, X'_{k+1})$ correspondant à la mesure de Poisson Q . Alors,

$$\begin{aligned} Z_{T_{k+1}}(dx, da) &= \tilde{Z}_{T_{k+1}}^k(dx, da) + \delta_{(X'_{k+1}, 0)}(dx, da) 1_{0 \leq \theta_{k+1} \leq m_1(T_{k+1}, \tilde{Z}_{T_{k+1}}^k, I_{k+1}, X'_{k+1})} \\ &\quad - \delta_{(X_{I_{k+1}}, A_{I_{k+1}}(\tilde{Z}_{T_{k+1}}^k))}(dx, da) 1_{m_1(T_{k+1}, \tilde{Z}_{T_{k+1}}^k, I_{k+1}, X'_{k+1}) \leq \theta_{k+1} \leq m_2(T_{k+1}, \tilde{Z}_{T_{k+1}}^k, I_{k+1}, X'_{k+1})} \\ &\quad + (\delta_{(X'_{k+1}, A_{I_{k+1}}(\tilde{Z}_{T_{k+1}}^k))}(dx, da) - \delta_{(X_{I_{k+1}}, A_{I_{k+1}}(\tilde{Z}_{T_{k+1}}^k))}(dx, da)) \\ &\quad \times 1_{m_2(T_{k+1}, \tilde{Z}_{T_{k+1}}^k, I_{k+1}, X'_{k+1}) \leq \theta_{k+1} \leq m_3(T_{k+1}, \tilde{Z}_{T_{k+1}}^k, I_{k+1}, X'_{k+1})} \end{aligned}$$

dépend aussi de façon unique de $(T_{k+1}, I_{k+1}, \theta_{k+1}, X'_{k+1}, \tilde{Z}^k, Z_{T_k})$ donc de (Z_0, Q) . Ceci prouve l'unicité forte de l'équation différentielle stochastique. \square

6.4.2 Propriétés de moments du processus

Nous nous intéressons à des propriétés de contrôle du processus en étudiant l'existence de moments du processus sur un intervalle de temps donné. En supposant que $\mathbb{E}(< Z_0, 1 >) < \infty$, il existe un unique processus vérifiant la définition précédente. La proposition suivante permet d'établir des propriétés de moments de ce processus.

Proposition 6.3. *Des propriétés de moments sur Z_0 engendrent des propriétés de moments sur le processus Z_t . Supposons qu'il existe $p \geq 1$ tel que $\mathbb{E}(< Z_0, 1 >^p) < +\infty$, alors pour tout $T < \infty$:*

$$\forall T > 0, \quad \mathbb{E} \left(\sup_{t \in [0, T]} < Z_t, 1 >^p \right) < +\infty.$$

La propriété de moment sera très utile pour les phénomènes limites. Pour démontrer cette propriété, il suffit de considérer un temps d'arrêt $\tau_N = \inf\{t \geq 0, \langle Z_t, 1 \rangle \geq N\}$ et de majorer $\sup_{t \in [0, T \wedge \tau_N]} \langle Z_t, 1 \rangle^p$ en négligeant les morts, de prendre l'espérance et d'utiliser un lemme de Gronwall. Ce procédé de démonstration est détaillé dans la thèse de V.C. Tran [39] ou l'article de S. Méléard [18] avec d'autres informations utiles.

Démonstration 6.3. On utilise à nouveau le temps d'arrêt $\tau_N = \inf\{t \geq 0, N_t \geq N\}$. On applique la formule de Itô et on majore en ne prenant en compte que les événements augmentant l'effectif total de la population c'est à dire les naissances uniquement :

$$\begin{aligned} \sup_{u \in [0, t \wedge \tau_N]} \langle Z_u, 1 \rangle^p &\leq \langle Z_0, 1 \rangle^p + \int_0^{t \wedge \tau_N} \int_{\tilde{\chi}} Q(ds, di, d\theta, dx') \\ &\quad \left[1_{\{0 \leq \theta < m_1(s, Z_{s-}, i, x')\}} ((\langle Z_{s-}, 1 \rangle + 1)^p - \langle Z_{s-}, 1 \rangle^p) \right]. \end{aligned}$$

Comme $(1+y)^p - y^p \leq C(p)(1+y^{p-1})$, en prenant l'espérance et en utilisant le théorème de Fubini qui s'applique grâce aux hypothèses énoncées et par le choix de τ_N , on obtient :

$$\begin{aligned} \mathbb{E} \left(\sup_{u \in [0, t \wedge \tau_N]} \langle Z_u, 1 \rangle^p \right) &\leq \mathbb{E}(\langle Z_0, 1 \rangle^p) \\ &\quad + \mathbb{E} \left(\int_0^{t \wedge \tau_N} C(p)(1 + \langle Z_s, 1 \rangle^{p-1}) \int_{\tilde{\chi}} b(x, a, Y_s^b) Z_s(dx, da) ds \right) \\ &\leq \mathbb{E}(\langle Z_0, 1 \rangle^p) + \bar{b}C(p) \int_0^t \mathbb{E}(\langle Z_{s \wedge \tau_N}, 1 \rangle + \langle Z_{s \wedge \tau_N}, 1 \rangle^p) ds. \end{aligned}$$

Puisque $\mathbb{E}(\langle Z_s, 1 \rangle + \langle Z_s, 1 \rangle^p) \leq 2\mathbb{E}(1 + \langle Z_s, 1 \rangle^p) \leq 2\mathbb{E} \left(1 + \sup_{u \in [0, s]} \langle Z_u, 1 \rangle^p \right)$,

on peut déduire que :

$$\mathbb{E} \left(\sup_{u \in [0, t \wedge \tau_N]} \langle Z_u, 1 \rangle^p \right) \leq \mathbb{E}(\langle Z_0, 1 \rangle^p) + 2\bar{b}C(p)T + 2\bar{b}C(p) \int_0^t \mathbb{E} \left(\sup_{u \in [0, s \wedge \tau_N]} \langle Z_u, 1 \rangle^p \right) ds.$$

Enfin, en utilisant l'inégalité de Gronwall, on obtient

$$\begin{aligned} \mathbb{E} \left(\sup_{u \in [0, t \wedge \tau_N]} \langle Z_u, 1 \rangle^p \right) &\leq (\mathbb{E}(\langle Z_0, 1 \rangle^p) + 2\bar{b}C(p)T) \exp[2\bar{b}C(p)T] \\ &\leq C(p, T), \end{aligned} \tag{6.12}$$

où $C(p, T)$ est une constante indépendante de N . En utilisant le fait que $\lim_{N \rightarrow +\infty} \tau_N = +\infty$ \mathbb{P} -p.s. et en appliquant le lemme de Fatou, on déduit le résultat voulu

$$\mathbb{E} \left(\sup_{u \in [0, t \wedge T]} \langle Z_u, 1 \rangle^p \right) \leq C(p, T) < \infty.$$

□

6.4.3 Études du générateur infinitésimal du processus Z_t

Afin de caractériser le processus d'évolution, nous analysons le générateur infinitésimal du processus Markovien $(Z_t)_{t \geq 0} \in \mathbb{D}(\mathbb{R}_+, \mathcal{M}_P(\tilde{\chi}))$. Il est complexe de caractériser directement le générateur du processus Z puisqu'il est à valeurs dans un espace mesure. Le générateur infinitésimal du processus $\langle Z_t, f_t \rangle$ est décrit pour un ensemble de fonctions tests f_t .

Les facteurs exogènes Y_t^b , Y_t^d et Y_t^e étant indépendants de la mesure de Poisson Q qui contrôle la dynamique du processus Z , nous considérons le générateur infinitésimal L_t^Y conditionné par les trajectoires des facteurs exogènes jusqu'à la date t .

Définissons $(\mathcal{Y}_t^{b,d,e})_{t \geq 0}$ la filtration associée aux facteurs Y^b , Y^d et Y^e . Le générateur infinitésimal L_t^Y est lui-même décrit sur un ensemble de fonctions Φ de sorte que $L_t^Y \Phi(\langle Z_t, f_t \rangle) = \frac{\partial}{\partial u} \mathbb{E} \left(\Phi(\langle Z_u, f_u \rangle) | \mathcal{Y}_u^{b,d,e} \right) |_{u=t}$. L'expression du générateur infinitésimal est donnée dans la proposition suivante :

Proposition 6.4. *Soit $Z_0 \in \mathcal{M}_P(\tilde{\chi})$ une mesure déterministe telle que :*

$$N_0 = \langle Z_0, 1 \rangle < +\infty.$$

Soit $\Phi \in \mathcal{C}_b(\mathbb{R})$ et $f : (x, a, s) \mapsto f_s(x, a) \in \mathcal{C}_b^{0,1,1}(\tilde{\chi} \times \mathbb{R}, \mathbb{R})$.

Le processus $\langle Z_t, f_t \rangle$ possède un générateur infinitésimal conditionnel défini par

$$\begin{aligned} L_t^Y \Phi(\langle Z_t, f_t \rangle) &= \int_{\tilde{\chi}} Z_t(dx, da) \left[\Phi'(\langle Z_t, f_t \rangle) \left(\frac{\partial f_t}{\partial a} + \frac{\partial f_t}{\partial s} \right)(x, a) \right. \\ &\quad + b(x, a, Y_t^b) \int_{\tilde{\chi}} [\Phi(\langle Z_t, f_t \rangle + f_t(x', 0)) - \Phi(\langle Z_t, f_t \rangle)] K^b(x, a, dx') \\ &\quad + d(x, a, Y_t^d) [\Phi(\langle Z_t, f_t \rangle - f_t(x, a)) - \Phi(\langle Z_t, f_t \rangle)] \\ &\quad \left. + e(x, a, Y_t^e) \int_{\tilde{\chi}} [\Phi(\langle Z_t, f_t \rangle + f_t(x', a)) - \Phi(\langle Z_t, f_t \rangle - f_t(x, a))] K^e(x, a, dx') \right]. \end{aligned} \quad (6.13)$$

Démonstration 6.4. Soient $\Phi \in \mathcal{C}_b(\mathbb{R})$ et $f : (x, a, s) \mapsto f_s(x, a) \in \mathcal{C}_b^{0,1,1}(\tilde{\chi} \times \mathbb{R}, \mathbb{R})$.

L'hypothèse de contrôle sur la variable Z_0 permet d'appliquer la Proposition 6.3 concernant l'existence de moments. Ainsi, on sait que pour tout $T > 0$:

$$\mathbb{E} \left(\sup_{u \in [0, T]} \langle Z_u, 1 \rangle \right) < C(T) < +\infty.$$

Soit $t > 0$. On rappelle que $L_t \Phi(\langle Z_t, f_t \rangle) = \frac{\partial}{\partial u} \mathbb{E} \left(\Phi(\langle Z_u, f_u \rangle) | \mathcal{Y}_u^{b,d,e} \right) |_{u=t}$.

En utilisant que $\Phi(\langle Z_u, f_u \rangle) = \Phi(\langle Z_0, f_0 \rangle) + \sum_{s \leq u} [\Phi(\langle Z_{s-} + (Z_s - Z_{s-}), f_s \rangle) - \Phi(\langle Z_{s-}, f_{s-} \rangle)]$

$\Phi(Z_{s-}, f_s >)]$, et en prenant l'espérance, on peut écrire :

$$\begin{aligned} \mathbb{E}[\Phi(< Z_u, f_u >)] &= \mathbb{E}[\Phi(< Z_0, f_0 >)] + \mathbb{E} \int_0^u \left(\Phi'(< Z_s, f_s >) < Z_s, \frac{\partial f_s}{\partial a} + \frac{\partial f_s}{\partial s} > \right. \\ &+ \mathbb{E} \left\{ \int_0^u \int_{\Xi} 1_{i \leq N_{s-}} \left([\Phi(< Z_s, f_s > + f_s(x', 0)) - \Phi(< Z_s, f_s >)] 1_{\{0 \leq \theta < m_1(s, Z_{s-}, i, x')\}} \right. \right. \\ &+ [\Phi(< Z_s, f_s > - f_s(x, a)) - \Phi(< Z_s, f_s >)] 1_{\{m_1(s, Z_{s-}, i, x') \leq \theta < m_2(s, Z_{s-}, i, x')\}} \\ &+ [\Phi(< Z_s, f_s > + f_s(x', a)) - \Phi(< Z_t, f_s > - f_s(x, a))] 1_{\{m_2(s, Z_{s-}, i, x') \leq \theta < m_3(s, Z_{s-}, i, x')\}} \Big\} \\ &\left. Q(ds, di, d\theta, dx') \right\} = \mathbb{E}[\Phi(< Z_0, f_0 >)] + \mathbb{E}[\Psi^Y(u, Z)], \end{aligned}$$

où

$$\begin{aligned} \Psi^Y(u, Z) &= \int_0^u \int_{\tilde{\chi}} ds Z_s(dx, da) \left[\Phi'(< Z_s, f_s >) \left(\frac{\partial f_s}{\partial a} + \frac{\partial f_s}{\partial s} \right)(x, a) \right. \\ &+ b(x, a, Y_s^b) \int_{\chi} [\Phi(< Z_s, f_s > + f_s(x', 0)) - \Phi(< Z_s, f_s >)] K^b(x, a, dx') \\ &+ d(x, a, Y_s^d) [\Phi(< Z_s, f_s > - f_s(x, a)) - \Phi(< Z_s, f_s >)] \\ &\left. + e(x, a, Y_s^e) \int_{\chi} [\Phi(< Z_s, f_s > + f_s(x', a)) - \Phi(< Z_t, f_s > - f_s(x, a))] K^e(x, a, dx') \right]. \end{aligned}$$

Soit $T > t > 0$, l'application $u \in [t, T] \mapsto \Psi^Y(u, Z)$ est :

– dominée par la fonction

$$T \left(\left(\left\| \frac{\partial f}{\partial a} \right\|_{\infty} + \left\| \frac{\partial f}{\partial s} \right\|_{\infty} \right) \|\Phi'\|_{\infty} + 2(\bar{b} + \bar{d} + \bar{e}) \|\Phi\|_{\infty} \right) \sup_{u \in [0, T]} < Z_u, 1 >$$

qui est intégrable d'après la propriété des moments.

– dérivable \mathbb{P} - presque sûrement en u de dérivée :

$$\begin{aligned} \frac{\partial \Psi^Y}{\partial u}(u, Z_u) &= \int_{\tilde{\chi}} Z_u(dx, da) \left[\Phi'(< Z_u, f_u >) \left(\frac{\partial f_u}{\partial a} + \frac{\partial f_u}{\partial u} \right)(x, a) \right. \\ &+ b(x, a, Y_u^b) \int_{\chi} [\Phi(< Z_u, f_u > + f_u(x', 0)) - \Phi(< Z_u, f_u >)] K^b(x, a, dx') \\ &+ d(x, a, Y_u^d) [\Phi(< Z_u, f_u > - f_u(x, a)) - \Phi(< Z_u, f_u >)] \\ &\left. + e(x, a, Y_u^e) \int_{\chi} [\Phi(< Z_u, f_u > + f_u(x', a)) - \Phi(< Z_t, f_u > - f_u(x, a))] K^e(x, a, dx') \right], \end{aligned}$$

– dominée par la fonction intégrable

$$\left(\left(\left\| \frac{\partial f}{\partial a} \right\|_{\infty} + \left\| \frac{\partial f}{\partial s} \right\|_{\infty} \right) \|\Phi'\|_{\infty} + 2(\bar{b} + \bar{d} + \bar{e}) \|\Phi\|_{\infty} \right) \sup_{u \in [0, T]} < Z_u, 1 >$$

Ainsi, en appliquant un théorème de dérivation sous l'espérance par rapport à u et en prenant $u = t$, on obtient l'expression (6.13) du générateur infinitésimal. \square

6.4.4 Dynamique du processus d'évolution

Afin de décrire la dynamique du processus d'évolution, nous présentons certaines propriétés martingales du processus Markovien $(Z_t)_{t \geq 0}$ qui seront très utiles dans la suite. Nous étudions plus précisément la dynamique de $\langle Z_t, f_t \rangle$ pour un ensemble de fonctions f_t en identifiant une partie martingale.

Proposition 6.5. *Soit $f : (x, a, s) \mapsto f_s(x, a) \in \mathcal{B}_b(\tilde{\chi} \times \mathbb{R}_+)$. Alors, $\forall t \in \mathbb{R}_+$,*

$$\begin{aligned} \langle Z_t, f_t \rangle &= \sum_{i=0}^{N_t} f_t(a_i(t), x_i(t)) \\ &= \langle Z_0, f_0 \rangle + \int_0^t ds \int_{\tilde{\chi}} Z_s(dx, da) \left[\left(\frac{\partial f_s}{\partial a} + \frac{\partial f_s}{\partial s} \right) (x, a) \right. \\ &\quad \left. + b(x, a, Y_s^b) \left(\int_{\chi} f_s(x', 0) k^b(x, a, x') P(dx') \right) - d(x, a, Y_s^d) f_s(x, a) \right. \\ &\quad \left. + e(x, a, Y_s^e) \left(\int_{\chi} [f_s(x', a) - f_s(x, a)] k^e(x, a, x') P(dx') \right) \right] \\ &\quad + M_t^f, \end{aligned}$$

avec M_t^f une martingale dont le crochet est donné par :

$$\begin{aligned} \langle M^f \rangle_t &= \int_0^t ds \int_{\tilde{\chi}} Z_s(dx, da) \left[b(x, a, Y_s^b) \int_{\chi} f_s^2(x', 0) k^b(x, a, x') P(dx') \right. \\ &\quad \left. + d(x, a, Y_s^d) f_s^2(x, a) + e(x, a, Y_s^e) \int_{\chi} \left(f_s(x', a) - f_s(x, a) \right)^2 k^e(x, a, x') P(dx') \right], \end{aligned}$$

et où les processus Y_s^b , Y_s^e et Y_s^d suivent la dynamique diffusive (6.4) ou la dynamique discrète (6.5)

Démonstration 6.5. La démonstration se fait en supposant cette forme, en appliquant la formule de Itô à $\langle Z_t, f_t \rangle^2$ et en analysant cette expression avec celle issue de la décomposition de Doob (unicité de la décomposition de Doob). En effet, en écrivant la formule de Itô sur $\langle Z_t, f_t \rangle^2$, on démontre que le processus

$$\begin{aligned} &\langle Z_t, f_t \rangle^2 - \langle Z_0, f_0 \rangle^2 - \int_0^t ds \int_{\tilde{\chi}} Z_s(dx, da) \left[\left(\frac{\partial f_s}{\partial a} + \frac{\partial f_s}{\partial s} \right) (x, a) \right. \\ &\quad \left. - 2 \langle Z_s, f_s \rangle \left(b(a, x, Y_s^b) \int_{\chi} f_s(x', 0) k^b(x, a, x') P(dx') + d(x, a, Y_s^d) f_s(x, a) \right. \right. \\ &\quad \left. \left. + e(a, x, Y_s^e) \int_{\chi} \left(f_s(x', a) - f_s(x, a) \right) k^e(x, a, x') P(dx') \right) \right] \\ &\quad - \langle M^f \rangle_t \end{aligned}$$

est une martingale nulle en 0.

En notant L_t^Y le générateur infinitésimal associé au processus Markovien Z_t donné par la Proposition 6.4, le processus $\Phi(< Z_t, f_t >) - \Phi(< Z_0, f_0 >) - \int_0^t L_s^Y \Phi(< Z_s, f_s >) ds$ est une martingale nulle.

Ainsi, le processus défini par :

$$\begin{aligned} & < Z_t, f_t >^2 - < Z_0, f_0 >^2 - \int_0^t ds \int_{\tilde{\chi}} Z_s(dx, da) \left[\left(\frac{\partial f_s}{\partial a} + \frac{\partial f_s}{\partial s} \right) (x, a) \right. \\ & - b(a, x, Y_s^b) \int_{\chi} \left((< Z_{s-}, f_s > + f_s(x', 0))^2 - < Z_{s-}, f_s >^2 \right) k^b(x, a, x') P(dx') \\ & + d(x, a, Y_s^d) \left((< Z_{s-}, f_s > - f_s(x, a))^2 - (< Z_{s-}, f_s >)^2 \right) \\ & \left. - e(a, x, Y_s^e) \int_{\chi} \left((< Z_{s-}, f_s > + f_s(x', a) - f_s(x, a))^2 - (< Z_{s-}, f_s >)^2 \right) k^e(x, a, x') P(dx') \right] \end{aligned}$$

est une martingale nulle en 0. En identifiant les deux compensateurs de $< Z_t, f_t >^2$ permettant d'obtenir la partie martingale, on obtient facilement l'expression voulue. \square

Cette décomposition de la dynamique de $< Z_t, f_t >$ avec une partie martingale caractérise le processus d'évolution de population. Elle sera d'ailleurs très utile pour établir le lien Micro/Macro.

6.4.5 Distribution moyenne de la population

Afin de décrire des propriétés de régularité de ce processus d'évolution, nous considérons le scénario moyen d'évolution. Nous analysons la distribution moyenne à un instant t de la population en âge et en traits et nous prouvons l'existence d'une densité par rapport à la mesure de Lebesgue.

On se place dans le cas où les traits n'évoluent pas dans le temps. La distribution moyenne de la population à l'instant t s'exprime par rapport à une fonction $f_t(x, a)$ où $f_t : \tilde{\chi} \times \mathbb{R}_+ \mapsto \mathbb{R}$. La densité $g_t(x, a)$ de la distribution moyenne $\mathbb{E}(< Z_t, f_t >)$ est donnée par :

$$\mathbb{E}(< Z_t, f_t >) = \int_{\tilde{\chi}} f_t(x, a) \mathbb{E} Z_t(dx, da) = \int_{\tilde{\chi}} f_t(x, a) g_t(x, a) P(dx) da.$$

On peut calculer $\mathbb{E}(< Z_t, f_t >)$

$$\begin{aligned} \mathbb{E}(< Z_t, f_t >) &= < \mathbb{E}(Z_0, f_0 >) + \int_0^t ds < \mathbb{E}(Z_s), \frac{\partial f_s}{\partial a} + \frac{\partial f_s}{\partial s} > \\ &+ \int_0^t ds \mathbb{E} \left(\int_{\tilde{\chi}} b(x, a, Y_s^b) Z_s(dx, da) \int_{\chi} f_s(x', 0) k^b(x, a, x') P(dx') \right) \\ &- \int_0^t ds \mathbb{E} \left(\int_{\tilde{\chi}} d(x, a, Y_s^d) Z_s(dx, da) f_s(x, a) \right). \end{aligned}$$

Afin de se débarrasser du premier terme d'évolution $\frac{\partial f_s}{\partial a} + \frac{\partial f_s}{\partial s}$, on effectue un changement de variable. On pose pour $t \geq 0$,

$$f_s(x, a) = \phi(x, a + t - s).$$

Cette écriture permet d'écrire que $\frac{\partial f_s}{\partial a} + \frac{\partial f_s}{\partial s} = 0$ ce qui va simplifier énormément l'équation précédente. En effet, on majore l'expression en enlevant le terme négatif issu de la mortalité et on obtient :

$$\begin{aligned} \mathbb{E}(\langle Z_t, f_t \rangle) &\leq \langle \mathbb{E}(Z_0, \phi(\cdot, \cdot + t)) \rangle \\ &\quad + \int_0^t ds \mathbb{E} \left(\int_{\tilde{\chi}} \int_{\chi} b(x, a, Y_s^b) Z_s(dx, da) k^b(x, a, x') \phi(x', t - s) P(dx') \right). \end{aligned}$$

Le premier terme possède une densité car on a supposé que la variable aléatoire Z_0 possédait une densité. Concernant le deuxième terme, on peut effectuer le changement de variable $\alpha = t - s$ et le théorème de Fubini permet d'invertir les intégrales $\int_{\tilde{\chi}}$ et \int_{χ} . Ainsi, on obtient

$$\begin{aligned} \mathbb{E}(\langle Z_t, f_t \rangle) &\leq \langle \mathbb{E}(Z_0, \phi(\cdot, \cdot + t)) \rangle \\ &\quad + \int_0^t \int_{\chi} \phi(x', \alpha) \mathbb{E} \left(\int_{\tilde{\chi}} b(x, a, Y_{t-\alpha}^b) Z_{t-\alpha}(dx, da) k^b(x, a, x') \right) P(dx') d\alpha. \end{aligned}$$

Si on pose $h(x', \alpha) = \mathbb{E} \left(\int_{\tilde{\chi}} b(x, a, Y_{t-\alpha}^b) Z_{t-\alpha}(dx, da) \right) k^b(x, a, x')$, elle correspond à une densité de probabilité jointe. On a donc majoré notre expression par une expression possédant une densité h , on peut donc en conclure que $\mathbb{E}(\langle Z_t, f_t \rangle)$ est une variable aléatoire à densité g_t car elle ne possède pas d'atomes. Il est cependant difficile de trouver une expression précise de cette densité g à cause du terme de mortalité pour lequel le changement de variable ne fonctionne pas.

Remarque 6.10. Une idée serait de considérer que la densité $Z_s(dx, da)ds$ peut s'écrire sous la forme $Z_s(dx, da)ds = q(ds)q_x(da)q_{x,a}(ds)$ et on pourrait éventuellement trouver une expression précise de cette densité g .

Si on devait interpréter la densité $g_t(x', \alpha)$, elle correspondrait à la densité des individus à l'instant $t - \alpha$ des individus d'âge a . On peut aussi voir que $t - \alpha$ correspond au temps de naissance des individus d'âge α au temps t . Cette expression permet de remarquer que l'évolution dans le futur de la population reste à densité en âge et en trait.

Rappelons la méthode classique pour démontrer l'existence d'une densité lorsque la population n'était pas structurée par âge [18]. En supposant que la population n'est pas structurée par âge et que les intensités d'évolution sont constantes en temps, il est possible

d'utiliser une autre idée. En effet, l'astuce précédente permettait de se débarrasser du terme $\frac{\partial f_s}{\partial a} + \frac{\partial f_s}{\partial s}$ mais si la population n'est pas structurée par âge, ce terme disparaît de lui même en prenant une fonction f indicatrice qui ne dépend pas du temps.

Supposons que l'ensemble des traits est à valeur dans \mathbb{R}^d et que la distribution K^b possède une densité k^b par rapport à la mesure de Lebesgue. Considérons un borélien A de $\mathbb{R}^+ \times \mathbb{R}^d$ de mesure de Lebesgue nulle. Pour tout $n \geq 1$, nous considérons le temps d'arrêt $\tau_n = \inf\{t \geq 0, \langle Z_t, 1 \rangle \geq n\}$. Ainsi, nous pouvons écrire pour tout $t \geq 0$ et pour tout $n \geq 1$,

$$\begin{aligned} \mathbb{E}(\langle Z_{t \wedge \tau_n}, 1_A \rangle) &= \mathbb{E}(\langle Z_0, 1_A \rangle) + \mathbb{E} \int_0^{t \wedge \tau_n} ds \int_{\mathcal{X}} Z_s(dx) b(x, Y_s^b) \int_{\mathcal{X}} 1_A(x') k^b(x, x') P(dx') \\ &\quad - \mathbb{E} \int_0^{t \wedge \tau_n} ds \int_{\mathcal{X}} Z_s(dx) d(x, Y_s^b) 1_A(x). \end{aligned}$$

Le premier terme est nul. Le deuxième terme est nul étant donné que la mesure P est supposée être absolument continue par rapport à la mesure de Lebesgue. Le troisième terme est négatif. On en déduit que $\mathbb{E}(\langle Z_{t \wedge \tau_n}, 1_A \rangle)$ est négatif et est donc nul. Or τ_n tend vers l'infini lorsque n tend vers l'infini ce qui permet de voir que $\mathbb{E}(\langle Z_t, 1_A \rangle) = 0$ sur tout borélien A de mesure nulle. On conclut que la distribution moyenne possède une densité par rapport à la mesure de Lebesgue. \square

Cependant, cette preuve n'est plus valable dans le cadre d'une population structurée par âge car on ne peut pas considérer directement que $\frac{\partial 1_A}{\partial a} = 0$ à moins d'ajouter de la régularité à frontière des boréliens A .

Remarque 6.11. *Nous avons traité le cas où les traits n'évoluaient pas dans le temps. Cette méthode du changement de variable ne s'appliquerait pas facilement lorsque les traits évoluent. En revanche, la deuxième méthode de S. Méléard et N. Fournier est facilement adaptable lorsque la population n'est pas structurée par âge. Lorsque la population est structurée par âge, il faudrait ajouter de la régularité à la frontière de sorte que " $\frac{\partial 1_A}{\partial a} = 0$ ".*

6.5 Renormalisation du processus microscopique

Nous abordons le cas des grandes populations pour lequel il existe une approximation du modèle microscopique d'évolution par un modèle macroscopique. Ce lien Micro/Macro va permettre de traduire un comportement intéressant et pertinent du modèle à grande échelle. Cette étude d'inspire des travaux de S. Méléard [18] et de la thèse de V.C. Tran [39].

Tout d'abord, nous renormalisons le processus dans le cadre de grandes populations et nous transcrivons les propriétés des moments et les propriétés de martingale au processus

renormalisé. Bien que stochastiques, les intensités d'évolution sont supposées bornées ce qui assure un contrôle suffisant. Ces propriétés de moments permettent d'étudier des comportements limites dans le cadre des grandes populations.

Afin d'étudier le processus renormalisé, nous définissons aussi $\mathcal{M}_P^n(E)$ comme l'ensemble des mesures renormalisées par n qui s'écrivent de la forme

$$\mathcal{M}_P^n(E) = \left\{ \frac{1}{n} m \mid m \in \mathcal{M}_P(E) \right\}.$$

6.5.1 Définitions et propriétés mathématiques

Nous considérons dans cette partie que le nombre d'individus est proportionnel à n et que chaque individu est renormalisé par un poids $\frac{1}{n}$. Nous supposons que les intensités d'évolution ne sont pas modifiées par la renormalisation. En effet, ces développements sont destinés à l'évolution d'une population d'individus et non l'évolution d'espèces biologiques éventuellement bactériennes. Il n'y a donc pas de raisons réelles de penser que les intensités de natalité et de mortalité changent d'échelle par renormalisation comme cela était éventuellement le cas dans d'autres études [18].

Supposons que $Z_0^n = \frac{1}{n} \sum_{i=1}^{N_0^n} \delta_{a_i(0), x_i(0)}$ converge vers $Z_0(dx, da) \in \mathcal{M}_{P(\tilde{\chi})}$ admettant une densité $m_0(x, a)$ par rapport à la mesure de référence $P(dx) \times da$ sur $\tilde{\chi}$.

Tout d'abord, nous rappelons les propriétés de moments et la propriété de martingale obtenus avec ces nouveaux processus renormalisés. Les preuves sont des généralisations directes des preuves précédentes. Ces deux propriétés permettent d'étudier les phénomènes de convergence de la suite $(Z^n)_{n \in \mathbb{N}}$ et de considérer des comportements limites.

Proposition 6.6. *Considérons la suite $(Z^n)_{n \in \mathbb{N}}$ définie précédemment. Supposons qu'il existe $p \geq 1$ tel que $\sup_{n \in \mathbb{N}^*} E(< Z_0^n, 1 >^p) < +\infty$, alors pour tout $0 < T < \infty$:*

$$\sup_{n \in \mathbb{N}^*} \mathbb{E} \left(\sup_{t \in [0, T]} < Z_t^n, 1 >^p \right) < +\infty.$$

Démonstration 6.6. La démonstration se déduit immédiatement de la propriété de moment pour le processus non renormalisé. Il faudra seulement considérer le temps d'arrêt renormalisé $\tau_N^n = \inf\{t \geq 0, < Z_t^n, 1 > \geq N\}$ à la place de τ_N . En effet, le fait de prendre un sup en n vient naturellement du fait qu'on obtient la même constante de majoration $C(p, T)$ qui ne dépend ni de N ni de n puisque les intensités d'évolution sont supposées inchangées par changement d'échelle. \square

Concernant la propriété de martingale du processus d'évolution, nous caractérisons aussi la dynamique de $\langle Z_t^n, f_t \rangle$ de la manière suivante :

Proposition 6.7. *Soit $f : (x, a, s) \mapsto f_s(x, a) \in \mathcal{B}_b(\tilde{\chi} \times \mathbb{R}_+)$. Alors, $\forall t \in \mathbb{R}_+$,*

$$\begin{aligned} \langle Z_t^n, f_t \rangle &= \sum_{i=0}^{N_t} f_t(a_i(t), x_i(t)) \\ &= \langle Z_0^n, f_0 \rangle + \int_0^t ds \int_{\tilde{\chi}} Z_s^n(dx, da) \left[\left(\frac{\partial f_s}{\partial a} + \frac{\partial f_s}{\partial s} \right) (x, a) \right. \\ &\quad \left. + b(x, a, Y_s^b) \left(\int_{\chi} f_s(x', 0) k^b(x, a, x') P(dx') \right) - d(x, a, Y_s^d) f_s(x, a) \right. \\ &\quad \left. + e(x, a, Y_s^e) \left(\int_{\chi} (f_s(x', a) - f_s(x, a)) k^e(x, a, x') P(dx') \right) \right] \\ &\quad + M_t^{n,f}, \end{aligned}$$

avec $M_t^{n,f}$ une martingale de carré intégrable issu de 0 et dont le crochet est donné par :

$$\begin{aligned} \langle M_t^{n,f} \rangle &= \frac{1}{n} \int_0^t ds \int_{\tilde{\chi}} Z_s^n(dx, da) \left[b(x, a, Y_s^b) \int_{\chi} f_s^2(x', 0) k^b(x, a, x') P(dx') \right. \\ &\quad \left. + d(x, a, Y_s^d) f_s^2(x, a) + e(x, a, Y_s^e) \int_{\chi} (f_s(x', a) - f_s(x, a))^2 k^e(x, a, x') P(dx') \right], \end{aligned} \quad (6.14)$$

et où les processus Y_s^b , Y_s^e et Y_s^d suivent la dynamique diffusive (6.4) ou la dynamique discrète (6.5).

Démonstration 6.7. Cette relation se prouve exactement de la même manière que la propriété de martingale du processus Z_t non renormalisé. On remarque que la renormalisation du processus Z en Z^n pourrait aussi s'interpréter en une renormalisation de la fonction f par la fonction $\frac{f}{n}$.

6.5.2 Étude du processus limite pour les grandes populations

Nous passons à l'étude des comportements limites en grandes populations. Nous allons voir que ces comportements limites permettent de prouver la convergence de cette suite de processus (Z^n) et que le processus limite ζ va suivre presque sûrement une équation aux dérivées partielles stochastique qui généralisera l'équation (6.1) aux dérivées partielles de McKendrick-Von Foerster.

Nous prouvons l'existence de solutions faibles mesures pour l'équation aux dérivées partielles ainsi qu'une unicité de la solution faible. Pour l'étude qui suit, nous effectuons les hypothèses suivantes :

Hypothèses 6.1. Nous supposons que le processus $(Z_0^n)_{n \in \mathbb{N}^*} \in \mathcal{M}_F(\tilde{\chi})$ converge en loi dans $(\mathcal{M}_F(\tilde{\chi}), \epsilon)$ vers une mesure $\zeta_0 \in \mathcal{M}_F(\tilde{\chi})$ éventuellement aléatoire.

Nous supposons aussi la condition sur les moments avec le fait qu'il existe $\eta \in]0, 1[$ tel que :

$$\sup_{n \in \mathbb{N}^*} \mathbb{E}(< Z_0^n, 1 >^{2+\eta}) < \infty.$$

Proposition 6.8. Soit $f : (x, a) \mapsto f(x, a) \in \mathcal{B}_b(\tilde{\chi})$. Sous les hypothèses précédentes, la suite (Z^n) définie précédemment converge en loi lorsque n tend vers l'infini vers un processus ζ_t . Avec une probabilité égale à 1, $< \zeta_t, f_t >$ vérifie pour tout $t \in \mathbb{R}_+$:

$$\begin{aligned} < \zeta_t, f_t > &= \int_{\tilde{\chi}} f_t(x, a) m_0(x, a) da dx + \int_0^t ds \int_{\tilde{\chi}} \zeta_s(dx, da) \left[\left(\frac{\partial f}{\partial a} + \frac{\partial f}{\partial s} \right) (x, a) \right. \\ &\quad + b(x, a, Y_s^b) \int_{\chi} f_s(x', 0) k^b(x, a, x') P(dx') - d(x, a, Y_s^d) f_s(x, a) \\ &\quad \left. + e(x, a, Y_s^e) \int_{\chi} (f_s(x', a) - f_s(x, a)) k^e(x, a, x') P(dx'), \right] \end{aligned} \quad (6.15)$$

où les facteurs Y^b , Y^d et Y^e vérifient les dynamiques discrètes (6.5) ou continues (6.4).

La démonstration repose sur un phénomène de tension et utilise plusieurs notions importantes comme le critère d'Aldous [1], le critère de Rebolledo [24] et le critère de Méléard-Roelly [32]. Ces travaux sont introduits dans l'article de S.Méléard et N. Fournier [18] et la thèse de V.C. Tran [39] et nous adaptons leur démarche à notre cadre d'évolution.

Démonstration 6.8. Nous rappelons les étapes de la démonstration avant de les détailler point par point. Afin de démontrer l'existence de la solution, nous cherchons à démontrer qu'il existe une valeur d'adhérence de la suite Z^n et qui vérifie l'équation (6.15). À la dernière étape nous démontrerons l'unicité de la solution à (6.15) ce qui prouvera la convergence de la suite. Nous allons adopter le schéma de démonstration suivant :

1. On démontre d'abord la tension de la suite Z^n lorsque $\mathcal{M}_F(\tilde{\chi})$ est munie de la topologie vague. Pour cela, il faut utiliser le critère de Aldous [1] et Rebolledo [24] et il suffit de démontrer la tension de la partie martingale et de la partie à variation finie de $< X_n, f >$ et de démontrer que :

$$\sup_n \mathbb{E} \left(\sup_{t \in [0, T]} |< Z_s^n, f >| \right) < \infty.$$

Ainsi, on démontre que (Z^n) converge vers un processus Z dans $\mathbb{D}([0, T], \mathcal{M}_F(\tilde{\chi}))$ où $\mathcal{M}_F(\tilde{\chi})$ est muni de la topologie vague.

2. Pour étendre cette convergence à la topologie faible, il faut utiliser un critère prouvé par S. Méléard et S. Roelly [32] : il suffit de démontrer que la suite $< Z^n, 1 >$ converge vers $< Z, 1 >$. On démontrera que le processus limite Z est continu par construction.

3. On identifie le processus limite Z et on prouve qu'il est l'unique solution de l'équation (6.15). On démontre que la partie martingale converge vers 0 et qu'il ne reste plus que la partie à variation finie. Dans le cadre étudié par Chi, la mesure ponctuelle ζ_t était déterministe. Ici, la partie martingale tend aussi vers 0 mais la mesure limite n'est pas déterministe puisque la partie à variation finie n'est pas déterministe. En effet, les intensités de mortalité et natalité dépendent de facteurs Y_t^d et Y_t^b et la partie à variation finie est stochastique.
4. Enfin, on démontre qu'il l'unicité de la solution à l'équation (6.15).

À présent, on rentre dans les détails afin de bien comprendre la démonstration point par point.

Afin de démontrer l'existence de la solution, on va montrer qu'il existe une valeur d'adhérence continue de la suite $(Z^n)_{n \in \mathbb{N}}$ qui vérifie l'équation (6.15).

1. On commence par établir la tension des lois de $(Z^n)_{n \in \mathbb{N}}$ considérées comme des mesures de probabilité sur $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(\tilde{\chi}), \nu))$ en munissant $\mathcal{M}_F(\tilde{\chi})$ de la topologie de la convergence vague.

L'idée est d'utiliser le critère de Roelly [32] (Théorème 2.1) qui dit qu'il est suffisant de prouver que pour toute fonction $f : (x, a) \mapsto f(x, a)$ appartenant à un ensemble dense de l'espace $\mathcal{C}_0(\tilde{\chi}, \mathbb{R})$ des applications continues de limite 0 lorsque $|a| \rightarrow +\infty$, la suite $Z^{n,f} = \langle Z^n, f \rangle$ est tendue sur $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$. On considère une fonction $f \in \mathcal{C}_0(\tilde{\chi}, \mathbb{R}) \cap \mathcal{C}_b^{0,1}(\tilde{\chi}, \mathbb{R})$ qui est dense sur $\mathcal{C}_0(\tilde{\chi}, \mathbb{R})$. Rappelons que la partie martingale $M_t^{n,f}$ a déjà été définie et que la partie à variation finie $V_t^{n,f}$ se déduit facilement par :

$$\begin{aligned} V_t^{n,f} &= Z_t^{n,f} - M_t^{n,f} = Z_0^{n,f} + \int_0^t ds \int_{\tilde{\chi}} Z_s^n(dx, da) \left[\frac{\partial f}{\partial a}(x, a) \right. \\ &\quad + b(x, a, Y_s^b) \int_{\chi} f(x', 0) k^b(x, a, x') P(dx') d(x, a, Y_s^d) f(x, a) \\ &\quad \left. + e(x, a, Y_s^e) \int_{\chi} (f(x', a) - f(x, a)) k^e(x, a, x') P(dx') \right]. \end{aligned} \quad (6.16)$$

En utilisant le critère d'Aldous [1] et Rebolledo [24], il suffit de montrer les propriétés suivantes :

- Soit \mathcal{T} dense dans \mathbb{R}_+ , pour tout t appartenant à \mathcal{T} , les lois de $(\langle M^{n,f} \rangle_t)_n \in \mathbb{N}^*$ et $(\langle M^{n,f} \rangle_t)_n \in \mathbb{N}^*$ sont tendues sur \mathbb{R} .
- $\forall T \geq 0, \forall \epsilon > 0, \forall \eta > 0, \forall \delta > 0, \forall n_0 \in \mathbb{N}$,

$$\begin{aligned} &\sup_{n \geq n_0} \mathbb{P}(|\langle M^{n,f} \rangle_{T_n} - \langle M^{n,f} \rangle_{S_n}| \geq \eta, T_n < S_n + \delta) \leq \epsilon \\ \text{et} \quad &\sup_{n \geq n_0} \mathbb{P}(|\langle V^{n,f} \rangle_{T_n} - \langle V^{n,f} \rangle_{S_n}| \geq \eta, T_n < S_n + \delta) \leq \epsilon \end{aligned}$$

, pour toute suite de couples de temps d'arrêt (S_n, T_n) de la filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ tels que $S_n \leq T_n \leq T$.

Pour démontrer le premier point, il suffit de prouver que :

$$\forall T > 0, \quad \sup_{n \in \mathbb{N}^*} \mathbb{E}(\sup_{t \in [0, T]} | < M^{n, f} >_t |) < +\infty \text{ et } \sup_{n \in \mathbb{N}^*} \mathbb{E}(\sup_{t \in [0, T]} | < V^{n, f} >_t |) < +\infty.$$

Avec l'expression (6.15) du crochet de $M^{n, f}$, on a

$$\sup_{n \in \mathbb{N}^*} \mathbb{E}(\sup_{t \in [0, T]} | < M^{n, f} >_t |) \leq \|f\|_\infty^2 T(\bar{b} + \bar{d} + 2\bar{e}) \sup_{n \in \mathbb{N}^*} \left(\frac{1}{n} \mathbb{E}(\sup_{t \in [0, T]} < Z_t^n, 1 >) \right),$$

qui est fini grâce aux propriétés de moments. De même, on peut écrire :

$$\sup_{n \in \mathbb{N}^*} \mathbb{E}(\sup_{t \in [0, T]} | < V^{n, f} >_t |) \leq T \left(\left\| \frac{\partial f}{\partial a} \right\|_\infty + \|f\|_\infty(\bar{b} + \bar{d} + 2\bar{e}) \right) \sup_{n \in \mathbb{N}^*} \mathbb{E} \left(\sup_{t \in [0, T]} < Z_t^n, 1 > \right),$$

qui est lui même fini par les mêmes arguments ce qui achève la démonstration du premier point.

Afin de prouver le deuxième point, on fixe $T \geq 0, \epsilon > 0, \eta > 0$ et on considère un couple de temps d'arrêt S_n et T_n tels que $0 \leq S_n \leq T_n \leq T$. On cherche δ et un rang n_0 tel que $T_n \leq S_n + \delta$ et que le deuxième point soit satisfait. On peut écrire pour le crochet de la partie martingale

$$\begin{aligned} \mathbb{E}(| < M^{n, f} >_{T_n} - < M^{n, f} >_{S_n} |) &\leq \mathbb{E} \left(\frac{1}{n} \int_{S_n}^{T_n} \|f\|_\infty^2 (\bar{b} + \bar{d} + 2\bar{e}) < Z_s^n, 1 > ds \right) \\ &\leq \frac{C \|f\|_\infty^2 \delta}{n} \mathbb{E} \left(\sup_{t \in [0, T]} (< Z_t^n, 1 >) \right) \leq \frac{C(T) \delta}{n} \end{aligned}$$

et la partie à variation finie

$$\mathbb{E}(| < V^{n, f} >_{T_n} - < V^{n, f} >_{S_n} |) \leq \delta \left(\left\| \frac{\partial f}{\partial a} \right\|_\infty + \|f\|_\infty(\bar{b} + \bar{d} + 2\bar{e}) \right) \sup_{n \in \mathbb{N}^*} \mathbb{E} \left(\sup_{t \in [0, T]} < Z_t^n, 1 > \right).$$

Il est donc facile de voir qu'il est possible de trouver δ et n_0 qui permet de satisfaire le deuxième point.

Remarquons qu'on a prouvé la tension de la suite $< Z^n, f >$ dans $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$. Par le même type d'arguments, on prouve aussi que la suite $< Z^n, 1 >$ est tendue dans $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$.

2. Maintenant que la tension en topologie vague est démontrée, on cherche à prouver une tension sous la topologie étroite. Pour cela, on utilise le théorème de Prokhorov qui permet d'extraire de $(Z^n)_{n \in \mathbb{N}}$ une sous suite convergente en loi dans $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(\tilde{\chi}), v))$. On note $(Z^{\Phi(n)})_{n \in \mathbb{N}^*}$ cette sous suite et Z un processus de loi la loi limite. Ayant aussi prouvé la tension de la suite $\langle Z^n, 1 \rangle_{n \in \mathbb{N}^*}$, quitte à extraire de nouveau, on peut supposer que $\langle Z^{\Phi(n)}, 1 \rangle_{n \in \mathbb{N}^*}$ converge en loi vers $\langle Z, 1 \rangle$ dans $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$. Comme par construction,

$$\sup_{t \in \mathbb{R}_+} \sup_{f \in \mathcal{C}_b^{0,1}(\tilde{\chi}, \mathbb{R})} |\langle Z_t^{\Phi(n)}, f \rangle - \langle Z_{t-}^{\Phi(n)}, f \rangle| \leq \frac{C \|f\|_\infty}{\Phi(n)}, \mathbb{P} - \text{presque sûrement},$$

le processus limite Z est presque sûrement continu, on peut utiliser le Théorème 3 de S. Méléard et S. Roelly [32] qui prouve que la sous suite $(Z^{\Phi(n)})_{n \in \mathbb{N}^*}$ converge également en loi dans $\mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(\tilde{\chi}), \acute{e}))$. Ainsi, on a prouvé la tension en topologie étroite.

3. Il faut maintenant identifier cette valeur d'adhérence en prouvant qu'elle satisfait l'expression (6.15), il faut démontrer que la partie martingale disparaît. On considère $f : (x, a, t) \mapsto f_t(x, a)$ dans $\mathcal{C}_b^{0,1,1}(\tilde{\chi})$ et il faut démontrer que la quantité suivante est nulle presque sûrement.

$$\begin{aligned} \Psi_t(Z) = & \langle Z_t, f_t \rangle - \langle Z_0, f_0 \rangle - \int_0^t \int_{\tilde{\chi}} Z_s(dx, da) ds \left[\left(\frac{\partial f_s}{\partial a} + \frac{\partial f_s}{\partial s} \right)(x, a) \right. \\ & + b(x, a, Y_s^b) \int_{\chi} f_s(x', 0) K^b(x, a, dx') - d(x, a, Y_s^d) f_s(x, a) \\ & \left. + e(x, a, Y_s^e) \int_{\chi} (f_s(x', a) - f_s(x', a)) K^e(x, a, dx') \right]. \end{aligned}$$

Soit $T \in \mathbb{R}$, afin de démontrer que $\mathbb{P} \left(\sup_{s \in [0, T]} |\Psi_t(Z)| \neq 0 \right) = 0$, on va démontrer que

$\mathbb{E} \left(\sup_{s \in [0, T]} \Psi_t(Z) \right) = 0$. Comme le processus $M_t^{f, \Phi(n)} = \Psi_t(Z^{\Phi(n)})$ est une martingale de crochet connu, on peut écrire :

$$\mathbb{E}(|M_t^{f, \Phi(n)}|)^2 \leq \mathbb{E}(|M_t^{f, \Phi(n)}|^2) = \mathbb{E}(\langle M_t^{f, \Phi(n)} \rangle) \leq \frac{tC}{\Phi(n)} \mathbb{E} \left(\sup_{s \in [0, t]} \langle Z_s^{\Phi(n)}, 1 \rangle \right).$$

Comme $\mathbb{E} \left(\sup_{s \in [0, t]} \langle Z_s^{\Phi(n)}, 1 \rangle \right)$ est majoré uniformément en n , on en déduit en utilisant l'inégalité de Doob que :

$$\forall t \in \mathbb{R}_+, \quad \lim_{n \rightarrow +\infty} \mathbb{E} \left(\sup_{t \in [0, T]} |M_t^{f, \Phi(n)}| \right) \leq 4 \lim_{n \rightarrow +\infty} \mathbb{E} \left(|M_T^{f, \Phi(n)}| \right) = 0.$$

Ainsi, pour montrer que $\forall t \in \mathbb{R}_+, \Psi_t(Z) = 0 \text{ } \mathbb{P} - p.s.$, il suffit de prouver que

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left(\sup_{t \in [0, T]} |M_t^{f, \Phi(n)}| \right) = \mathbb{E} \left(\sup_{t \in [0, T]} \Psi_t(Z) \right).$$

Par conséquent, il faut prouver une propriété d'uniforme intégrabilité de la variable $\sup_{t \in [0, T]} \Psi_t(Z^{\Phi(n)})$ pour pouvoir prendre la limite sous l'espérance. Pour cela, en utilisant que $\sup_{t \in [0, T]} \Psi_t$ est continue et que f est continue bornée à dérivées continues

bornées, on majore la fonction $\sup_{t \in [0, T]} \Phi_t(z)$. Ainsi, pour tout $T \in \mathbb{R}_+$ et pour tout $z \in \mathbb{D}(\mathbb{R}_+, (\mathcal{M}_F(\tilde{\chi})))$, on a

$$\sup_{t \in [0, T]} |\Psi_t(z)| \leq C(T, f) \sup_{s \in [0, T]} (\langle z_s, 1 \rangle) := \Theta_T(z).$$

Par la propriété des moments, $\Theta_T(Z^{\Phi(n)})$ est uniformément intégrable (Proposition 2.2, p.494 de [16]) alors il en est de même pour $\sup_{t \in [0, T]} \Psi_t(Z^{\Phi(n)})$ ce qui prouve le résultat. Ainsi, la valeur d'adhérence vérifie bien cette équation.

4. La dernière étape consiste à démontrer l'unicité de la solution. Supposons qu'il existe deux solutions ζ_t^1 et ζ_t^2 de (6.15) issues de ζ_0 et tels que \mathbb{P} - presque sûrement, on a :

$$\forall t \in \mathbb{R}_+, \quad \zeta_t^1(dx, da) = \zeta_t^2(dx, da).$$

Tout d'abord, on peut remarquer les solution de (6.15) sont toutes presque sûrement continues (au sens L^1 faible) et restent de masse finie en temps fini :

$$\begin{aligned} \forall t \geq 0, \langle \zeta_t, 1 \rangle &= \langle \zeta_0, 1 \rangle + \int_0^t ds \int_{\tilde{\chi}} \zeta_s(dx, da) [b(x, a, Y_s^b) - d(x, a, Y_s^d)] \\ &\leq \langle \zeta_0, 1 \rangle + \int_0^t \bar{b} \langle \zeta_s, 1 \rangle ds \\ &\leq \langle \zeta_0, 1 \rangle \exp(\bar{b}t) < +\infty \text{ } \mathbb{P} - p.s. \end{aligned}$$

En considérant une fonction $\Phi \in \mathcal{C}_b^{0,1}(\tilde{\chi})$ telle que $\|\Phi\|_\infty \leq 1$, on peut définir une fonction $f \in \mathcal{C}_b^{0,1}(\tilde{\chi} \times \mathbb{R})$ telle que $f(x, a, s) = \Phi(x, a + t - s)$.

On peut écrire pour $i \in \{1, 2\}$, \mathbb{P} - presque sûrement, et pour tout $t \in [0, T]$:

$$\begin{aligned} \langle \zeta_t^i, \Phi \rangle &= \int_{\tilde{\chi}} \Phi(x, a + t) \zeta_0^i(dx, da) \\ &\quad + \int_0^t ds \int_{\tilde{\chi}} \zeta_s^i(dx, da) \left[b(x, a, Y_s^b) \int_{\chi} \Phi(x', t - s) K^b(x, a, dx') \right. \\ &\quad \left. - d(x, a, Y_s^d) \Phi(x, a + t - s) \right. \\ &\quad \left. + e(x, a, Y_s^e) \int_{\chi} (\Phi(x', a + t - s) - \Phi(x, a + t - s)) K^e(x, a, dx') \right]. \end{aligned}$$

La norme en variation totale sur l'espace $\mathcal{M}_F(\tilde{\chi})$ est définie pour μ et ν par :

$$\|\mu - \nu\|_{TV} = \sup\{|\langle \mu, f \rangle - \langle \nu, f \rangle|, f \in \mathcal{B}_b(\tilde{\chi}), \|f\|_\infty \leq 1\}$$

et V.C. Tran a démontré dans [39] que :

$$\|\mu - \nu\|_{TV} = \sup\{|\langle \mu, f \rangle - \langle \nu, f \rangle|, f \in \mathcal{C}_b(\tilde{\chi}, \mathbb{R}), \|f\|_\infty \leq 1\}.$$

Ainsi, \mathbb{P} - presque sûrement, pour tout $t \in [0, T]$, on a alors :

$$\begin{aligned} |\langle \zeta_t^1 - \zeta_t^2, \Phi \rangle| &\leq \int_0^t ds \left| \int_{\tilde{\chi}} (\zeta_s^1(dx, da) - \zeta_s^2(dx, da)) \left[d(x, a, Y_s^d) \Phi(x, a + t - s) \right. \right. \\ &\quad + b(x, a, Y_s^b) \int_{\chi} \Phi(x', t - s) k^b(x, a, x') P(dx') \\ &\quad \left. \left. + e(x, a, Y_s^e) \int_{\chi} (\Phi(x', a + t - s) - \Phi(x, a + t - s)) k^e(x, a, x') P(dx') \right] \right| \\ &\leq (\bar{b} + \bar{d} + 2\bar{e}) \int_0^t \|\zeta_s^1 - \zeta_s^2\|_{TV} ds. \end{aligned}$$

Ainsi, en prenant le sup en Φ dans le membre de gauche, en remarquant que tout fonction $\Phi \in \mathcal{C}_b(\tilde{\chi})$ est limite simple bornée d'une suite de fonctions de $\mathcal{C}_b^{0,1}(\tilde{\chi})$, on prouve que \mathbb{P} - presque sûrement

$$\forall t \in [0, T], \quad \|\zeta_t^1 - \zeta_t^2\|_{TV} \leq (\bar{b} + \bar{d}) \int_0^t \|\zeta_s^1 - \zeta_s^2\|_{TV} ds.$$

Le lemme de Gronwall assure que $\|\zeta_t^1 - \zeta_t^2\|_{TV} = 0$. Le temps T étant arbitraire, les mesures ζ_1 et ζ_2 coïncident \mathbb{P} - presque sûrement.

□

Remarque 6.12. *On peut noter le fait qu'en l'absence de facteurs dans les intensités d'évolution, la mesure limite était déterministe. Ici, la présence des facteurs permet d'avoir un processus mesure limite ζ qui vérifie une relation presque sûrement.*

6.5.3 Approximation par un modèle macroscopique aléatoire

Nous avons vu que le processus limite ζ_t était caractérisé comme étant l'unique solution de l'équation (6.15). Nous allons analyser la forme faible de cette équation ce qui nous fournira l'équation aux dérivées partielles stochastique vérifiée presque sûrement par la densité de cette mesure. Comme dans le cadre du processus de naissance et mort, il est possible de définir $g(\omega, x, a, t)$ la densité "aléatoire" de la mesure ponctuelle ζ_t .

Proposition 6.9. *La densité $g(\omega, a, x, t)$ du processus ζ_t est presque sûrement l'unique solution au sens classique (avec les propriétés de régularité suffisantes) du système suivant :*

$$\begin{aligned} \left(\frac{\partial g}{\partial t} + \frac{\partial g}{\partial a} \right) (\omega, x, a, t) &= - \left[d(x, a, Y_t^d(\omega)) + e(x, a, Y_t^e(\omega)) \right] g(\omega, x, a, t) \\ &\quad + \int_{\chi} e(x', a, Y_t^e) k^e(x', a, x) g(\omega, x', a, t) P(dx') \\ g(\omega, x, 0, t) &= \int_{\tilde{\chi}} b(x', a, Y_t^b(\omega)) k^b(x', a, x) g(\omega, x', a, t) P(dx') da \\ g(\omega, x, a, 0) &= g_0(\omega, x, a), \end{aligned} \tag{6.17}$$

avec les facteurs Y^b , Y^d et Y^e les solutions fortes uniques de leur équation de diffusion de type (6.4)

Démonstration 6.9. Puisque le processus ζ_t est solution presque sûre de l'équation (6.15), on en déduit que g est presque sûrement la solution faible du système d'équations (6.17). On cherche à passer d'une solution faible à une solution au sens classique. Afin de prouver que $g(\omega, a, x, t)$ n'est pas seulement une solution faible mais aussi une solution au sens standard, on a besoin de régularité sur la fonction g ce qui n'est pas simple à prouver directement.

On rappelle que les paramètres μ et σ , décrivant la dynamique des facteurs des intensités d'évolution, sont donnés de manière à assurer l'unicité forte de la dynamique de ces facteurs (caractère lipschitzien dans le cas d'une dynamique continue). Ainsi, on fixe une trajectoire des intensités d'évolution et l'équation (6.17) correspond à la généralisation de l'équation de McKendrick-Von Foerster par âge et par traits, avec les traits évoluant dans le temps. Il a été démontré par Perthame [34] que sous certaines hypothèses, qui sont largement remplies dans ce cadre, cette équation possédait une unique solution au sens classique avec les propriétés de régularité suffisantes. On en déduit que g est l'unique solution du système précédent au sens classique. \square

Cette équation permet de comprendre comment la population évolue dans le temps. Tout d'abord, il existe de nombreux points communs avec l'équation de McKendrick-Von Foerster classique. La population diminue à cause des décès des individus apparaissant dans le temps au travers du terme $-d(x, a, Y_t^d(\omega)) g(\omega, x, a, t)$.

La condition aux bord $g(\omega, x, 0, t) = \int_{\tilde{\chi}} b(x', a, Y_t^b(\omega)) k^b(x', a, x) g(\omega, x', a, t) P(dx') da$ (à l'âge nul) caractérise la naissance des individus et est assez complexe puisqu'à une date donnée, les naissances dépendent de toute la structure de la population.

Cependant, cette équation généralise l'équation (6.1) de McKendrick-Von Foerster par deux aspects. D'une part, la présence d'une évolution de traits justifie la présence du terme négatif $-e(x, a, Y_t^e(\omega)) g(\omega, x, a, t)$ et du terme positif $\int_{\chi} e(x', a, Y_t^e) k^e(x', a, x) g(\omega, x', a, t) dx'$.

Le premier terme retire des individus et le deuxième terme remplace ces individus par de nouveaux individus dont les traits sont tirés dans la distribution k^e . À un instant donné et à un âge donné, la densité sur un trait x va donc dépendre de la densité sur tous les autres traits ce qui complexifie l'équation d'évolution.

D'autre part, l'équation aux dérivées partielles est stochastique du fait de la présence des facteurs aléatoires dans la dynamique des intensités d'évolution. Elle permet de tenir compte de l'évolution des intensités d'évolution dans le temps et de leur aléa.

Remarque 6.13. *Dans le cas où il n'y a pas d'évolution des traits dans le temps, la densité $g(\omega, a, x, t)$ est presque sûrement l'unique solution au sens classique du système suivant :*

$$\begin{aligned} \left(\frac{\partial g}{\partial t} + \frac{\partial g}{\partial a} \right) (\omega, x, a, t) &= -d \left(x, a, Y_t^d(\omega) \right) g(\omega, x, a, t) \\ g(\omega, x, 0, t) &= \int_{\tilde{\chi}} b \left(x', a, Y_t^b(\omega) \right) k^b(x', a, x) g(\omega, x', a, t) P(dx') da \\ g(\omega, x, a, 0) &= g_0(\omega, x, a). \end{aligned}$$

Cette équation est une généralisation aléatoire de l'équation (6.1) de McKendrick-Von Foerster dont on sait qu'elle possède une unique solution [39, 34].

Remarque 6.14. *Nous aurions pu facilement tenir compte d'une éventuelle interaction dans le modèle [39], mais nous n'avons pas jugé que cela soit pertinent dans la mesure où la compétition au sein d'espèce humaine est difficile à modéliser précisément. Si nous en avons tenu compte, nous aurions obtenu une généralisation de l'équation logistique de type Gurtin-McCamy dite non linéaire et qui est donc plus difficile à traiter que cette équation qui est linéaire.*

6.5.4 Interprétation de l'équation aux dérivées partielles stochastique

La densité du processus limite est solution presque sûre d'une équation aux dérivées partielles stochastique. Cette équation généralise l'équation (6.1) de McKendrick-Von Foerster qui est classique en dynamique de population ce qui confirme le fait que le modèle possède un comportement macroscopique cohérent en généralisant ces modèles macroscopiques classiques. En effet, en utilisant le modèle microscopique, on cherche généralement à fournir des informations macroscopiques sur la structure de la population. Ce résultat assure que le modèle établit un lien micro/macro pertinent et que les informations macroscopiques seront intéressantes à étudier.

Le fait d'avoir une équation macroscopique stochastique confirme le caractère doublement aléatoire de cette modélisation. On sait qu'un modèle microscopique est aléatoire par nature mais généralement, lorsqu'on passe aux approximations macroscopiques, son comportement

devient généralement déterministe [39]. Ici, le comportement macroscopique reste aléatoire et il faudra étudier l'impact de cet aléa dans l'étude pratique d'une dynamique de population. Cette équation aux dérivées partielles stochastique étant complexe, il est plus commode d'utiliser une simulation du modèle à l'échelle microscopique même dans le cadre d'une population très grande.

Dans le cas où la population est simplement structurée en ne considérant pas la présence des facteurs exogènes, cette équation macroscopique se ramène à l'équation de McKendrick Von Foerster :

$$\begin{aligned} \left(\frac{\partial g}{\partial t} + \frac{\partial g}{\partial a} \right) (a, t) &= -d(a)g(a, t) \\ g(0, t) &= \int_{\tilde{\chi}} b(a)g(a, t)da \\ g(a, 0) &= g_0(a). \end{aligned} \tag{6.18}$$

L'étude d'une solution d'équilibre est assez classique [43] et revient à chercher une solution stationnaire $g(a)$. L'étude de l'équation fournit la solution :

$$g(a) = g(0)e^{-\int_0^a d(\alpha) d\alpha}.$$

En injectant la solution dans la condition au bord en âge nul, on obtient :

$$g(0) = g(0) \int_0^{+\infty} b(a)e^{-\int_0^a d(\alpha) d\alpha} da.$$

Ainsi, la condition de l'existence d'un équilibre stable dans ce cas s'écrit :

$$R_0 = \int_0^{+\infty} b(a)e^{-\int_0^a d(\alpha) d\alpha} da = 1.$$

La variable R_0 est appelé taux de reproduction net. Il est possible de voir que si $R_0 < 1$, il y a extinction exponentielle et que si $R_0 > 1$, il y a explosion exponentielle [43]. On s'intéresse à l'étude d'un équilibre dans le cas général de notre modèle en étendant la notion de taux de reproduction net.

6.5.5 Taux de reproduction net

Le taux de reproduction net R_0 correspond au nombre moyen d'enfants au cours de sa vie. Cette notion caractérise le comportement d'évolution générationnelle de la population. En effet, il faut qu'un individu ait en moyenne plus d'un enfant au cours de sa vie afin d'assurer le renouvellement des générations. On calcule ce taux R_0 pour une population asexuée structurée par âge uniquement et pour laquelle les intensités d'évolution ne dépendent pas du temps. L'expression classique du taux R_0 est donnée par :

$$R_0 = \int_0^{+\infty} b(a)e^{-\int_0^a d(\alpha) d\alpha} da.$$

Ce taux de reproduction correspond à l'intégrale du taux de naissance pondérée par la probabilité de survivre jusqu'à l'âge a . La position de R_0 par rapport à 1 va déterminer l'évolution future de la population. En effet, si $R_0 < 1$, le seul équilibre stable est l'extinction de la population. Si $R_0 = 1$, on est dans le cas où la population converge vers un équilibre. Enfin, si $R_0 > 1$, on peut montrer qu'il existe un unique équilibre "asymptotiquement exponentiellement stable" [43] et sa densité asymptotique par âge $g(a)$ est donnée par :

$$g(a) = \frac{\lambda e^{\lambda a} \pi(0, a)}{\eta \int_0^\infty e^{-\lambda \alpha} \pi(0, \alpha) d\alpha},$$

où $\pi(a_1, a_2) = S_0(0, a_1, a_1 + a_2)$ correspond à la probabilité qu'un individu d'âge a_1 vive jusque l'âge a_2 (voir Section 6.3.4) et où λ est choisi tel que $1 = \int_0^\infty e^{-\lambda a} b(a) \pi(0, a) da$.

Dans ce cadre, on définit plus généralement un taux de reproduction net $R_{t_0}(x, a)$ comme le nombre moyen d'enfants qu'un individu d'âge a et de trait x à la date t_0 aura au cours de sa vie. Cette notion permet de généraliser la notion standard de taux de reproduction net en tenant compte de l'impact des traits et du temps. En effet, les individus sont aussi caractérisés par des traits qui peuvent évoluer dans le temps et les intensités d'évolution dépendent de facteurs. Dans ce cadre, on suggère la définition suivante du taux $R_{t_0}(x, a)$:

$$\begin{aligned} \tilde{X}_t(x, a) &= (X_t^{x,a}, A_t^{x,a}), \\ R_{t_0}(X_{t_0}^{x,a}, A_{t_0}^{x,a}) &= \mathbb{E} \left(\int_{t_0}^{+\infty} b(X_t^{x,a}, A_t^{x,a}, Y_t^b) e^{-\int_{t_0}^t d(X_u^{x,a}, A_u^{x,a}, Y_u^d) du} dt \middle| \mathcal{F}_{t_0} \right), \end{aligned}$$

où le processus markovien $\tilde{X}_t^{x,a}$ décrit l'évolution dans le temps de l'âge et des traits d'un individu d'âge a et de trait x en date 0. Le processus $X_t^{x,a}$ est chaîne de Markov inhomogène en temps décrivant l'évolution des caractéristiques à l'aide de l'intensité e de changement de trait. Ainsi, pour l'étude d'évolution d'une population structurée par âge et par traits, on pourra étudier ce nouveau taux de reproduction net et analyser sa position par rapport à 1.

6.6 Conclusion

Nous avons présenté dans ce chapitre un modèle microscopique de dynamique de population qui est adapté à l'étude d'une population humaine. En effet, il tient compte de nombreux phénomènes d'évolution (naissance, décès, évolution des traits, émigration, immigration) et permet de considérer des intensités d'évolution assez générales pouvant dépendre de facteurs aléatoires dans le temps. Nous avons décrit une méthode de simulation intéressante de ce modèle permettant de déduire des scénarios d'évolution.

De plus, nous avons étudié le modèle dans le cadre d'une grande population. Nous avons établi un lien micro/macro qui souligne le fait que le comportement du modèle à l'échelle

macroscopique est pertinent. Le modèle peut être approximé par un modèle macroscopique aléatoire qui généralise le modèle classique de McKendrick-Von Foerster et le modèle de Perthame par âge et par traits [34].

Nous verrons dans les deux chapitres suivants les applications pratiques de ce modèle de dynamique de populations dans des sujets variés. Dans le chapitre 7, nous aborderons les projections démographiques de la population française dans le futur ainsi que l'étude du problème des retraites. Dans le chapitre 8, nous étudierons l'aspect financier du risque de longévité dans le domaine de l'assurance-vie. Nous verrons que ce modèle fournit des scénarios pertinents dont l'analyse permet de recueillir des informations intéressantes.

6.7 Annexes

Annexe A : Renormalisation pour la simulation d'un processus

On considère un événement se produisant dans à des temps poissonniens au taux $\lambda(s)$ à la date s . Cela revient à étudier la simulation de :

$$\int_0^t \int_{\mathbb{R}_+} 1_{\theta \leq \lambda(s)} Q(ds, d\theta)$$

où $Q(ds, d\theta)$ est une mesure de Poisson sur $\mathbb{R}_+ \times \mathbb{R}_+$ d'intensité $dsd\theta$. On suppose un caractère borné du taux à savoir que pour tout s , $\lambda(s) \leq \bar{\lambda}$. Il est possible de considérer le "Thinning" du processus de Poisson associé à cette contrainte s'écrivant $Q'(ds, d\theta) = 1_{\lambda(s) \leq \bar{\lambda}} Q(ds, d\theta)$ sur $\mathbb{R}_+ \times [0, \bar{\lambda}]$ dont l'intensité peut s'écrire $\bar{\lambda} ds d\theta / \bar{\lambda}$.

$$\int_0^t \int_{\mathbb{R}_+} 1_{\theta \leq \lambda(s)} Q'(ds, d\theta)$$

La simulation de ce processus revient à simuler une variable exponentielle E de paramètre $\bar{\lambda}$ déterminant le temps d'occurrence et de simuler une variable uniforme U sur $[0, 1]$ représentant $d\theta / \bar{\lambda}$. Si $U < \lambda(s) / \bar{\lambda}$ il y a effectivement occurrence de l'événement sinon il ne se passe rien.

Annexe B : Étude d'une population structurée par âge uniquement

Se restreindre au cas d'une population uniquement structurée par âge permet de simplifier l'ensemble des résultats et permet de retrouver des résultats plus classiques. Cette étude sera très utile en pratique (dans le domaine de la démographie) dans la mesure où il est rare d'avoir des informations précises sur les traits d'une population et sur la distribution des traits à la naissance. De plus, les études classiques en démographie considèrent un taux de mortalité et un taux de naissance dépendant uniquement de l'âge et du temps. Il n'y a

plus d'évolution de traits dans ce cadre.

La définition du processus d'évolution de population est plus simple et plus intuitif. Nous rappelons certains résultats déjà exposés dans le cas général et nous ajoutons quelques commentaires. Nous redéfinissons le processus d'évolution comme suit :

$$\begin{aligned} Z_t &= \sum_{i=0}^{N_0} \delta_{(A_i(Z_0)(t,0,A_i(Z_0)))} \\ &+ \int_0^t \int_{\tilde{\chi}} 1_{i \leq N_{s-}} [\delta_{t-s} 1_{0 \leq \theta < m_1(s, Z_{s-}, i)} \\ &- \delta_{A_i(Z_{s-}) + t - s} 1_{m_1(s, Z_{s-}, i) \leq \theta < m_2(s, Z_{s-}, i)}] Q(ds, di, d\theta), \end{aligned}$$

où $A_i(Z_{s-})$ correspond à l'âge d'un individu i de la population à la date s et où

$$\begin{aligned} m_1(s, Z_{s-}, i) &= b(A_i(Z_{s-}), Y_s^b). \\ m_2(s, Z_{s-}, i) &= m_1(s, Z_{s-}, i) + d(A_i(Z_{s-}), Y_s^d). \end{aligned}$$

La méthode de simulation reste la même mais toutes les naissances sont équivalentes et génèrent un enfant caractérisé uniquement par son âge (0 à la naissance). Nous présentons un corollaire de la proposition 5.

Corollaire 6.1. *Soit $f : (a, s) \mapsto f_s(a) \in \mathcal{B}_b(\mathbb{R}_+ \times \mathbb{R}_+)$. Alors, $\forall t \in \mathbb{R}_+$,*

$$\begin{aligned} \langle Z_t, f_t \rangle &= \sum_{i=0}^{N_t} f_t(a_i(t)) \\ &= \langle Z_0, f_0 \rangle + \int_0^t ds \langle Z_s, \frac{\partial f_s}{\partial a} + \frac{\partial f_s}{\partial s} \rangle \\ &+ \int_0^t ds \int_{\mathbb{R}_+} Z_s(da) [b(a, Y_s^b) f_s(0) - f_s(a) d(a, Y_s^d)] ds \\ &+ M_t^f, \end{aligned}$$

avec M_t^f une martingale dont le crochet est donné par :

$$\langle M^f \rangle_t = \int_0^t ds \int_{\mathbb{R}_+} Z_s(da) [b(a, Y_s^b) f_s^2(0) + f_s^2(a) d(a, Y_s^d)] ds,$$

et où les processus de diffusion Y_s^d et Y_s^b suivent la dynamique (6.4)

Étude de la distribution moyenne

On démontre que dans le cadre d'une population uniquement structurée par âge, la distribution moyenne possède une densité en âge. En effet, on peut écrire pour tout $t \geq 0$:

$$\begin{aligned}\mathbb{E}(\langle Z_t, f_t \rangle) &= \int_{\tilde{\mathcal{X}}} f_t(a) \mathbb{E}(Z_t(da)) \\ &= \langle \mathbb{E}(Z_0, f_0) \rangle + \int_0^t ds \langle \mathbb{E}(Z_s), \frac{\partial f_s}{\partial a} + \frac{\partial f_s}{\partial s} \rangle \\ &\quad + \int_0^t f_s(0) \mathbb{E} \left(\int_{\mathbb{R}_+} b(a, Y_s^b) Z_s(da) \right) \\ &\quad - \int_0^t \int_{\mathbb{R}_+} f_s(a) \mathbb{E} \left(\int_{\mathbb{R}_+} d(a, Y_s^d) E(Z_s(da)) \right) ds.\end{aligned}$$

On pose $\phi(a + t - s) = f_s^t(a)$, ce qui nous donne $\frac{\partial f_s}{\partial a} + \frac{\partial f_s}{\partial s} = 0$. En supposant que la population initiale Z_0 possède une densité m_0 (une pyramide des âges), on peut écrire

$$\begin{aligned}\mathbb{E}(\langle Z_t, f_t \rangle) &= \int_{\mathbb{R}_+} \phi(a + t) m_0(a) da + \int_0^t \phi(t - s) \mathbb{E} \left(\int_{\mathbb{R}_+} b(a, Y_s^b) Z_s(da) \right) ds \\ &\quad - \int_0^t \int_{\mathbb{R}_+} \phi(a + t - s) \mathbb{E} \left(\int_{\mathbb{R}_+} d(a, Y_s^d) E(Z_s(da)) \right) ds.\end{aligned}$$

En majorant cette expression en ne tenant plus compte du terme négatif de mortalité, on obtient :

$$\mathbb{E}(\langle Z_t, f_t \rangle) \leq \int_{\mathbb{R}_+} \phi(a + t) m_0(a) da + \int_0^t \phi(t - s) \mathbb{E} \left(\int_{\mathbb{R}_+} b(a, Y_s^b) Z_s(da) \right) ds.$$

On effectue le changement de variable $\alpha = a + t$ dans la première et $\alpha = t - s$ dans la deuxième, on peut écrire :

$$0 \leq \mathbb{E}(\langle Z_t, f_t \rangle) \leq \int_{\mathbb{R}_+} \phi(\alpha) m_0(\alpha - t) d\alpha + \int_0^t \phi(\alpha) \mathbb{E} \left(\int_{\mathbb{R}_+} b(a, Y_{t-\alpha}^b) Z_{t-\alpha}(da) \right) d\alpha.$$

Ainsi, la mesure associée à $\mathbb{E}(Z_t)$ est bornée par une mesure à densité sans atome. On en déduit que Z_t ne possède pas d'atome et est une mesure absolument continue par rapport à la mesure de Lebesgue. On ne peut pas calculer explicitement sa densité à cause du terme en mortalité qui complexifie la forme de la densité. Le changement de variable $\alpha = a + t - s$ ne permet pas de se ramener au cas d'une variable à densité.

Estimation de la densité

Finalement, $\mathbb{E}(Z_t)$ possède une densité en âge ce qui permettra de dire que la moyenne de nos simulation de population doit posséder une pyramide des âges à tout instant. On

peut estimer cette densité avec les simulations. En effet, si on a effectué N simulation, la moyenne empirique donnée par :

$$M_t^N(da) = \frac{1}{N} \sum_{i=1}^N Z_t^i(da) \sim \mathbb{E}(Z_t)(da).$$

L'approximation de la densité se fait en convolant cette densité estimée par la densité gaussienne $\kappa_h(x) = \frac{1}{\sqrt{2\pi h}} e^{-\frac{x^2}{2h}}$. On pose $\hat{g}_t^N(a) = \kappa(h) \times M_t^N = \frac{1}{N} \sum_{i=1}^N \kappa_h(\alpha - a) Z_t^i(da)$. La loi forte des grands nombres donne lorsque N tend vers l'infini :

$$\lim_{N \rightarrow \infty} \hat{g}_t^N(a) = \int_{\mathbb{R}_+} \kappa_h(\alpha - a) E(Z_t(d\alpha)).$$

Il est connu que :

$$g_t(a) = \lim_{h \rightarrow 0} \int_{\mathbb{R}_+} \kappa_h(\alpha - a) E(Z_t(d\alpha)).$$

Ainsi, il est possible de faire une estimation non paramétrique de la densité par âge de la population moyenne.

Remarque 6.15. *L'estimation de la densité se fait par une méthode de Monte Carlo. On peut noter que cela revient aussi à se placer dans le cadre d'une très grande population en effectuant une renormalisation. Cela fournit l'intuition que le cadre des grandes populations permet de régulariser le processus et d'obtenir la densité de la distribution moyenne.*

Cas des grandes populations

Concernant le cadre des grandes populations, on peut réécrire l'étude menée dans le cas général. Le processus limite vérifie l'équation simplifiée suivante :

$$\begin{aligned} \langle \zeta_t, f \rangle &= \int_{\mathbb{R}_+} f(a) m_0(a) da + \int_0^t ds \int_{\mathbb{R}_+} \zeta(da) \frac{\partial f}{\partial a} \\ &+ \int_0^t ds \int_{\mathbb{R}_+} \zeta(da) [b(a, Y_s^b) f(0) - f(a) d(x, a, Y_s^d)] ds, \end{aligned}$$

où les facteurs Y^b et Y^d ont une dynamique donnée. Cette forme faible permet de déduire une équation aux dérivées partielles stochastique. En effet, la densité aléatoire g vérifie avec probabilité 1 le système suivant :

$$\begin{aligned} \left(\frac{\partial g_t}{\partial t} + \frac{\partial g_t}{\partial a} \right) (\omega, a) &= -d(a, Y_t^d) g_t(\omega, a) \\ g_t(\omega, 0) &= \int_0^\infty b(a, Y_t^b) g_t(\omega, a) da \\ Y_t^b &= Y_0^b + \int_0^t \mu^b(s, Y_s^b) ds + \int_0^t \sigma^b(s, Y_s^b) dB_s^b \\ Y_t^d &= Y_0^d + \int_0^t \mu^d(s, Y_s^d) ds + \int_0^t \sigma^d(s, Y_s^d) dB_s^d. \end{aligned}$$

Cette équation correspond à l'équation (6.1) de McKendrick-Von Foerster qui admet une unique solution régulière classique. On en déduit que la densité aléatoire est l'unique solution classique suivant et possède les bonnes propriétés de régularité.

Bibliographie

- [1] Aldous. D. (1978), “Stopping times and tightness”, *Annals of Probability*, 6, 335-340.
- [2] Antolin P. and H. Blommestein(2007), “Governments and the market for longevity-indexed bonds”, *OECD Journal : Financial Market Trends*, 4, 92.
- [3] Athreya K.B. and P.E. Ney (1970), “Branching Processes.”, Springer edition.
- [4] Bellman R. and T.E. Harris (1952), “On age-dependent binary branching processes”, *Annals of Mathematics*, 55, 280-295.
- [5] Bensusan H. and N. El Karoui (2009), “Microscopics Models for Populations Trait and Age Structures and Applications for Longevity problem”, Preprint.
- [6] Bolker B. and S. Pacala (1997), “Using moment equations to understand stochastically driven spatial pattern formation in ecological systems”, *Theoretical population biology* 52, 179-197.
- [7] Blake D., A. Cairns and K. Dowd (2006), “A two-factor model for stochastic mortality with parameter uncertainty : Theory and calibration”, *Journal of Risk and Insurance*, 73, 687-718
- [8] Caswell H. (2001), “Matrix Population Models, Construction, Analysis, and Interpretation”, Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts.
- [9] Champagnat N. (2004), “Mathematical study of stochastic models of evolution belonging to the ecological theory of adaptive dynamics”, Dissertation.
- [10] Charlesworth N. (1994), “Evolution in Age-structured Population”, *Cambridge University Press*, 2 edition.
- [11] Crump K.S. and C.J. Mode (1968), “A general age-dependent branching process i”, *J. Math. Anal. Appl.*, 24, 494-508.
- [12] Crump K.S. and C.J. Mode (1969), “A general age-dependent branching process ii”, *J. Math. Anal. Appl.*, 25, 8-17.
- [13] Desvillettes L., R. Ferrière, and C. Prévost (2004), “Infinite dimensional reaction-diffusion for population dynamics”, *Preprint*, CMLA ENS Cachan.
- [14] Doney R.A. (1972), “Age-dependent birth and death processes”, *Z. Wahrscheinlichkeitstheorie verw.*, 22, 69-90.

- [15] Dieckman U., B. Ernande and M. Heino (2004), “Adaptive changes in harvested populations : plasticity and evolution of age and size at maturation”, *Proc. R. Soc. Lond. B*, 271, 415-423.
- [16] Ethier S.N. and T.G. Kurtz (1986), “Markov Processus, Characterization and Convergence”, *John Wiley & Sons, New York*.
- [17] Ferrière R. and V.C. Tran (2009), “Stochastic and deterministic models for age-structured populations with genetically variable traits”, *ESAIM : Proceedings*, Vol 27, 289-310.
- [18] Fournier N. and S. Méléard (2004), “A microscopic probabilistic description of a locally regulated population and macroscopic approximations”, *Annals of Applied Probability*, 14, 4, 1880-1919.
- [19] Galton F. and H.W. Watson (1874), “On the probability of the extinction of families”, *J. Anthropol. Inst. Great B. and Ireland*, Vol 4, 138-144.
- [20] Harris T.E. (1963), “The Theory of Branching Processes”, *Springer, Berlin*.
- [21] Imhoff E.V. and W. Post (1998), “Microsimulation methods for population projection”, *Population*, 1, 97-136.
- [22] Jagers P. (1969), “A general stochastic model for population development”, *Skand. Aktuarietidskr*, 52, 84-103.
- [23] Jagers P. (1975), “Branching Processes with Biological Applications”, *Wiley edition*.
- [24] Joffe A. and M. Métivier (1986), “Weak convergence of sequences of semimartingales with applications to multitype branching processes”, *Advances in Applied Probability*, 18, 20-65.
- [25] Kendall D.G. (1949), “Stochastic processes and population growth”, *J. Roy. Statist. Sec., Ser. B*, 11, 230-264.
- [26] Lambert A. (2007), “Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct”, *Annals of Applied Probability*, 7, 1, 134-155.
- [27] Law R. and U. Dieckmann (2000), “Moment approximations of individual-based models in The Geometry of Ecological Interactions”, *Cambridge University press*.
- [28] Lee R.D. and L.R. Carter (1992), “Modeling and forecasting U.S. mortality”, *Journal of the American Statistical Association*, 87, 659-675.
- [29] Lotka A.J. (1922), “The stability of the normal age distribution”, *Proc. Nat. Acad. Science*, Vol 8, 339-345.
- [30] Méléard S. and V.C. Tran (2009), “Trait substitution sequence process and canonical equation for age-structured populations”, *Journal of Mathematical Biology*, Vol 58, No 6, 881-921.
- [31] McKendrick A.G (1926), “Applications of mathematics to medical problems”, *Proc. Edin. Math.Soc.*, Vol 54, 98-130.

- [32] Méléard S. and S. Roelly (1993), “Sur les convergences étroite ou vague de processus à valeurs mesures”, *C.R.Acad.Sci.Paris*, Serie I, 317, 785-788.
- [33] Mischler S., B. Perthame and L. Ryzhik (2002), “Stability in a non linear population maturation model”, *Mathematical Models and Methods in Applied Sciences*, Vol 12, No 12, 1-22.
- [34] Perthame B. and L. Ryzhik (2005), “Exponential decay for the fragmentation or cell-division equation”, *Journal of the Differential Equations*, 210, 155-177.
- [35] Oelschläger. K. (1990), “Limit theorem for age-structured populations”, *The Annals of Probability*, 18, 1, 290-318.
- [36] Rotenberg. M. (1983), “Transport theory for growing cell populations”, *Journal of Theoretical Biology*, 103, 181-199.
- [37] Sharpe F.R. and A.J. Lotka (1911), “A problem in age distribution”, *Philos. Mag.*, Vol 21, 435-438.
- [38] Solomon W. (1987), “Representation and approximation of large population age distributions using Poisson random measures”, *Stochastic Processes and their Applications*, 26, 237-255.
- [39] Tran V.C. (2006), “Modèles particuliers stochastiques pour des problèmes d’évolution adaptative et pour l’approximation de solutions statistiques”, Dissertation.
- [40] Tran V.C. (2008), “Large population limit and time behaviour of a stochastic particle model describing an age-structured population”, *ESAIM : P & S*, Vol 12, 345-386.
- [41] Von Foerster H., (1959), “Some remarks on changing populations”, *The Kinetics of Cellular Proliferation*, In Grune & Stratton, editor, New York, 382-407.
- [42] Wang F.J.S., (1977), “A central limit theorem for age- and density-dependent population processes”, *Stochastic Processes and their Applications*, 5, 173-193.
- [43] Webb G.F., (1985), “Theory of Nonlinear Age-dependent population dynamics”, *Monographs and Textbooks in Pure and Applied mathematics*, Vol 89.

BIBLIOGRAPHIE

Chapitre 7

Projections démographiques et étude du problème des retraites

Introduction

Dans ce chapitre, nous nous intéressons aux projections démographiques de la population française ainsi qu'à une étude du problème des retraites. Le vieillissement spectaculaire des populations des pays occidentaux nous invite à modéliser au mieux la dynamique d'une population d'individus.

Motivations

Le vieillissement de la population des pays occidentaux résulte d'un taux de fécondité actuel faible, et d'une amélioration nette des conditions de vie [25, 27, 23]. Le débat concernant la longévité des individus et l'évolution future de la population est très controversé dans la mesure où les estimations données par les démographes se révèlent souvent assez éloignées de ce qui se passe en réalité. D'un point de vue probabiliste, on peut dire que l'évolution des taux démographiques n'est pas déterministe mais aléatoire. Il est par conséquent impossible de prédire à l'avance l'évolution de la situation démographique dans le futur. L'idée est donc de modéliser la dynamique de la population afin d'obtenir un scénario moyen d'évolution ainsi qu'un panel de scénarios possibles appartenant à un intervalle de confiance donné. Les scénarios en dehors de cet intervalle de confiance sont des scénarios "extrêmes" qui fournissent de l'information dont il faut tenir compte pour dresser un bilan démographique.

De plus, le vieillissement de la population s'accompagne de problématiques économiques importantes. Un article d'Antolin et Blommestein [1] montre que l'amélioration de la longévité des personnes de plus de 80 ans a un impact important sur le PIB d'un pays ainsi

que sur les décisions politiques des gouvernements. Le paiement des retraites fait partie des problématiques majeures liées au vieillissement de la population. Le système français des retraites est un système de répartition et de solidarité intergénérationnelle dans lequel les actifs paient la retraite des inactifs. Le gouvernement qui assure le paiement des retraites doit donc s'assurer de maintenir une situation démographique "acceptable" pour que le financement des retraites dans le futur soit possible.

Résumé de l'étude

Dans ce chapitre, nous introduisons tout d'abord les modèles utilisés pour mener une étude démographique. Nous considérons le modèle de mortalité détaillé dans le chapitre 5 de la thèse ainsi que le modèle microscopique de dynamique de population détaillé dans le chapitre 6 de la thèse. Nous incorporons le modèle du taux de mortalité au sein du modèle de dynamique de population. Cette modélisation tient compte de phénomènes démographiques importants comme l'évolution de l'âge et des caractéristiques des individus, les naissances, les décès, l'occurrence d'éventuelles pandémies guerres ou catastrophes naturelles ainsi que les flux migratoires. Ce modèle microscopique possède des propriétés intéressantes dont celle de fournir un lien "micro/macro" : à partir d'informations sur chacun des individus, nous pouvons déduire une évolution "cohérente" d'une population et en tirer des informations à l'échelle macroscopique.

Après avoir analysé la situation démographique de la population française en 2007, nous allons utiliser le modèle afin d'obtenir des scénarios d'évolution de la population. Nous analysons plus spécifiquement l'évolution de la taille de la population et de la répartition des individus par âge. Dans une première étude, nous tenons compte d'une immigration et d'une natalité statiques et nous ne prenons pas en compte les traits des individus et leur évolution. Le scénario moyen d'évolution est assez "cohérent" avec les prévisions des démographes et un panel de scénarios permet d'avoir une idée des évolutions possibles de la structure de la population. L'analyse des différents scénarios donnent une information plus riche que le scénario moyen avec par exemple un âge maximal de la pyramide des âges qui est différent. La stochasticité du modèle est double dans la mesure où il y a non seulement l'aléa des taux démographiques mais aussi l'aléa du processus d'évolution de population. À très long terme, le panel de scénarios devient de plus en plus large avec des scénarios extrêmes de plus en plus extrêmes. Nous verrons aussi l'importance des scénarios des taux démographiques sur l'évolution de la population.

Ensuite, nous nous intéressons au problème des retraites et nous analysons d'un point de vue démographique les solutions envisageables. Les deux solutions proposées d'une immigration choisie et d'une réforme sur l'âge de départ à la retraite sont actuellement sources de nombreux débats politiques. Nous allons adopter un point de vue démographique plu-

tôt qu'économique et nous allons voir si ces solutions peuvent modifier la structure de la population. En effet, le système français des retraites est un système par répartition et par solidarité intergénérationnelle où les actifs paient pour les inactifs. La notion de *potential support ratio* ou *dependency ratio* correspondant au rapport du nombre d'actifs sur le nombre d'inactifs est un facteur démographique important concernant le paiement des retraites. Le niveau actuel de ce ratio est à 4 et les démographes considèrent qu'un *potential support ratio* de 3 dans le futur est un niveau raisonnable économiquement. Nous nous intéressons au problème des retraites sur une durée de 50 ans et l'objectif démographique sera donc de maintenir un *potential support ratio* supérieur à 3 jusqu'en 2050.

De nombreuses études estiment que l'immigration dans les pays développés pourrait être la solution au problème du paiement des retraites dans le futur. Nous verrons que cette stratégie seule n'est pas suffisante ne permettant pas à d'avoir une situation démographique raisonnable jusqu'en 2050. En effet, le nombre nécessaire d'immigrés est totalement incompatible avec la politique actuelle des gouvernements d'Europe. Nous cherchons à combiner une réforme sur l'âge de départ à la retraite à une politique d'immigration choisie afin de proposer une politique envisageable par le gouvernement. Nous proposons une réforme de départ à la retraite associée à une politique d'immigration raisonnable permettant de satisfaire cet objectif démographique.

7.1 Cadre d'évolution d'une population française

Dans cette partie, nous mettons en place la modélisation de la dynamique d'une population d'êtres humains ainsi que la modélisation des taux démographiques. Les modèles utilisés sont rappelés et adaptés au cadre de l'évolution d'une population d'individus.

7.1.1 Modèle de dynamique de population

Nous rappelons le modèle microscopique d'évolution d'une population détaillé au chapitre 6 de cette thèse qui s'inspire de l'approche de N. Fournier et S. Méléard [15] et de V.C. Tran [28, 18].

Processus d'évolution

Nous proposons une modélisation microscopique dans laquelle chaque individu est caractérisé par son âge $a \in \mathbb{R}_+$ et son trait $x \in \chi$. χ est un espace de trait qui décrit l'ensemble des caractéristiques des individus. Chaque individu est une particule représentée par une masse de Dirac. La structure de la population est donc une mesure ponctuelle constituée d'un ensemble de masses de Dirac. Notons $\mathcal{M}_P(E)$ l'ensemble des mesures ponctuelles sur E .

Les évènements affectant la population sont les naissances, les décès et l'évolution de traits. Les occurrences sont modélisées à l'aide d'une mesure ponctuelle de Poisson. Entre deux évènements, les individus vieillissent continûment avec le temps.

Afin de décrire le processus d'évolution de la population qui généralise le processus BPLD, nous définissons sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ les éléments aléatoires indépendants suivants :

- La population initiale $Z_0 \in \mathcal{M}_P(\tilde{\chi})$ de taille finie est une mesure ponctuelle telle que $\mathbb{E}(\langle Z_0, 1 \rangle) < \infty$.
- $Q(ds, di, d\theta, dx')$ est une mesure ponctuelle de Poisson sur l'espace $\mathbb{R}_+ \times \Xi = \mathbb{R}_+ \times \mathbb{N}^* \times \mathbb{R}_+ \times \chi$ dont le compensateur est $ds \, n(di) \, d\theta \, P(dx')$. Dans cette notation, ds est lié au temps, $n(di) = \sum_{k \geq 1} \delta_k(di)$ est une mesure ponctuelle de comptage sur \mathbb{N}^* pour sélectionner un individu, $d\theta$ est un artifice afin de décrire l'occurrence des évènements et $P(dx)$ la mesure de référence sur l'espace des traits χ .
- Les facteurs exogènes Y_t^b , Y_t^d et Y_t^e sont des facteurs communs à l'ensemble de la population (à l'échelle de la population et non pas à l'échelle de l'individu), qui sont indépendants de Q , de Z_0 et qui sont indépendants entre eux.

Notons $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ la filtration canonique engendrée par Z_0 , Q et par les facteurs Y^b , Y^d et Y^e . Le processus Z à valeurs dans $\mathcal{M}_P(\tilde{\chi})$ est caractérisé par l'évolution suivante :

$$\begin{aligned} Z_t = & \sum_{i=0}^{N_0} \delta_{(X_i(Z_0), A_i(Z_0)+t)} + \int_0^t \int_{\Xi} 1_{i \leq N_{s-}} \left[\delta_{(x', t-s)} 1_{0 \leq \theta < m_1(s, Z_{s-}, i, x')} \right. \\ & - \delta_{(X_i(Z_{s-}), A_i(Z_{s-})+t-s)} 1_{m_1(s, Z_{s-}, i, x') \leq \theta < m_2(s, Z_{s-}, i, x')} \\ & \left. + (\delta_{(x', A_i(Z_{s-})+t-s)} - \delta_{(X_i(Z_{s-}), A_i(Z_{s-})+t-s)}) 1_{m_2(s, Z_{s-}, i, x') \leq \theta < m_3(s, Z_{s-}, i, x')} \right] Q(ds, di, d\theta, dx'), \end{aligned}$$

où $X_i(Z_t)$ et $A_i(Z_t)$ correspondent respectivement aux traits et à l'âge de l'individu i de la population à la date t et où :

$$\begin{aligned} m_1(s, Z_{s-}, i, x') &= b(X_i(Z_{s-}), A_i(Z_{s-}), Y_s^b) k^b(X_i(Z_{s-}), A_i(Z_{s-}), x'). \\ m_2(s, Z_{s-}, i, x') &= m_1(s, Z_{s-}, i, x') + d(X_i(Z_{s-}), A_i(Z_{s-}), Y_s^d) k^d(X_i(Z_{s-}), A_i(Z_{s-}), x'). \\ m_3(s, Z_{s-}, i, x') &= m_2(s, Z_{s-}, i, x') + e(X_i(Z_{s-}), A_i(Z_{s-}), Y_s^e) k^e(X_i(Z_{s-}), A_i(Z_{s-}), x'). \end{aligned}$$

$b(x, a, Y_t^b)$ représente l'intensité de natalité qui dépend des caractéristiques de l'individu (son âge a et ses traits x) ainsi que d'un facteur exogène Y_t^b qui commun l'ensemble de la population. Ce facteur exogène, éventuellement stochastique, caractérise l'évolution de l'intensité dans le temps. De même, l'intensité de mortalité est représentée par l'expression

$d(x, a, Y_t^d)$ et l'intensité d'évolution des traits est représentée par $e(x, a, Y_t^e)$.

Remarque 7.1. *L'ensemble des propriétés mathématiques, l'ensemble des hypothèses sur les intensités démographiques ainsi que la méthode de simulation du processus sont décrits au chapitre 6 de la thèse.*

Lien Micro/Macro

Contrairement à de nombreux modèles de dynamique de population qui sont des modèles macroscopiques, ce modèle est un modèle microscopique qui décrit l'évolution d'une population à l'échelle de l'individu. La modélisation microscopique présente de nombreux avantages même si elle est tout de même plus complexe à mettre en œuvre [16].

La modélisation microscopique permet de tenir compte d'informations précises sur chaque individu et fournit des informations plus riches en sortie. De plus, la modélisation est aléatoire et l'analyse des scénarios donne lieu à des interprétations intéressantes. Il faudra faire attention au fait que l'on considère un échantillon fictif de la population totale et non la population entière. Cela peut introduire un biais puisque l'échantillon fictif peut avoir une structure légèrement différente de la structure réelle la population.

De plus, cette modélisation microscopique possède l'avantage majeur de faire un lien micro/macro pertinent. Lorsque la taille de la population est grande, le modèle microscopique peut être approximé à grande échelle par un modèle macroscopique. En effet, lorsque la taille de la population tend vers l'infini, le processus d'évolution converge vers un processus solution presque sûre d'une équation aux dérivées partielles stochastiques. En effet, nous avons démontré dans le chapitre 6 que la densité $g(\omega, x, a, t)$ de la distribution de la population par âge et par traits vérifie :

$$\begin{aligned} \left(\frac{\partial g}{\partial t} + \frac{\partial g}{\partial a} \right) (\omega, x, a, t) &= - \left[d(x, a, Y_t^d(\omega)) + e(x, a, Y_t^e(\omega)) \right] g(\omega, x, a, t) \\ &\quad + \int_{\mathcal{X}} e(x', a, Y_t^e) k^e(x', a, x) g(\omega, x', a, t) P(dx') \\ g(\omega, x, 0, t) &= \int_{\mathcal{X} \times \mathbb{R}_+} b(x', a, Y_t^b(\omega)) k^b(x', a, x) g(\omega, x', a, t) P(dx') da \\ g(\omega, x, a, 0) &= g_0(\omega, x, a). \end{aligned} \tag{7.1}$$

Cette équation généralise l'équation classique en démographie de McKendrick Von-Foerster [30, 31]. Par conséquent, ce modèle microscopique possède de bonnes propriétés macroscopiques et peut être utilisé pour l'étude de l'évolution de grandes populations.

7.1.2 Taux de mortalité

La mortalité des pays occidentaux a diminué de manière spectaculaire au cours du vingtième siècle [19]. L'évolution de la mortalité est considérée comme stochastique dans la mesure où il n'est pas possible de la prédire avec certitude. D'ailleurs, la croissance de la population est due en grande partie à la baisse de la mortalité infantile [7, 8].

L'intensité de mortalité d du modèle de dynamique de population est modélisée afin de caractériser au mieux l'évolution dans le temps de la mortalité des individus. En pratique, nous modélisons la mortalité en décrivant son taux de mortalité q . Le taux de mortalité sera déterminé par un modèle de Cairns à deux facteurs par âge. Il existe un lien entre l'intensité de mortalité d et le taux de mortalité q .

Modélisation d'un taux de mortalité individuel

De nombreuses caractéristiques individuelles ont un impact important sur le niveau de mortalité comme la catégorie sociale, le niveau d'éducation, le statut matrimonial pour n'en citer que quelques unes [20]. L'impact de ces caractéristiques a été étudié en détails dans le chapitre 5. Nous rappelons la modélisation stochastique du taux individuel de mortalité.

Notons $\mathcal{A} = \{A_l, 1 \leq l \leq N\}$ la partition des classes d'âge et N le nombre de tranches d'âge. Le taux de mortalité $q(y, x, a, t)$ correspond à la probabilité de décès à la date t d'un individu de traits $x = (x_i)_{1 \leq i \leq M_x}$ d'âge a vivant dans un environnement (ou une région) de caractéristiques $y = (y_i)_{1 \leq i \leq M_y}$.

En fait, cette régression logistique sur chaque tranche d'âge peut s'écrire comme suit.
 $\forall l \in \{1, \dots, N\}, \forall a \in A_l,$

$$\begin{aligned} \text{logit}[q(y, x, a, t)] &= \log \left(\frac{q(y, x, a, t)}{1 - q(y, x, a, t)} \right) \\ &= \alpha_1(t) + a\alpha_2(t) + \sum_{i=1}^{M_x} \beta_i^l(t)^\top x_i + \sum_{j=1}^{M_y} \gamma_j^l(t)^\top y_j. \end{aligned}$$

L'étude de ce modèle, la signification précise de chacune de ces variables et de ces paramètres ainsi que la calibration des paramètres du modèle sont détaillées dans le chapitre 5. Rappelons tout de même l'interprétation des différentes variables.

- Les variables y_i correspondent aux variables qui décrivent les caractéristiques globales.
- Les variables x_i correspondent aux variables qui décrivent les caractéristiques individuelles explicatives (les traits) d'un individu i .
- $N_{x,i}$ correspond au nombre de classes au sein de la variable individuelle x_i et $N_{y,i}$ correspond au nombre de classes au sein de la variable globale $y(i)$. Les variables peuvent être quantitatives avec une seule classe ou catégorielles avec un certain nombre de classes.

- L'entier N correspond au nombre de tranches d'âge dans le modèle.
- Le coefficient $\alpha_1(t)$ correspond au facteur d'amélioration naturelle de longévité et le facteur $\alpha_2(t)$ correspond à l'amélioration spécifique à un âge donné.
- Les coefficients $(\beta_i^l)_{\substack{1 \leq i \leq M_x \\ 1 \leq l \leq N}}$ et $(\gamma_i^l)_{\substack{1 \leq i \leq M_y \\ 1 \leq l \leq N}}$ correspondent aux coefficients de la régression logistique, les premiers étant associés aux traits individuels et les autres aux variables globales d'environnement.

Remarque 7.2. *Nous pouvons noter le caractère borné du taux de mortalité q . La modélisation se fait sur le logit de mortalité qui peut appartenir à \mathbb{R} mais le taux de mortalité q se retrouve à l'aide de la fonction inverse du logit*

$$q = \text{logit}^{-1}(\text{logit}(q)) = \frac{1}{1 + \exp(-\text{logit}(q))} \in [0, 1].$$

Prise en compte des événements catastrophiques

Nous proposons aussi de prendre en compte des catastrophes naturelles ou pandémies éventuelles qui modifient la structure de la population. Les événements exceptionnels peuvent être pris en compte en les considérant comme une composante supplémentaire du taux de mortalité comme cela est fait dans les études biologiques ou certaines études démographiques. Par conséquent, il faut estimer le taux d'occurrence de ces événements rares à l'aide d'un historique.

En analysant les courbes de mortalité, on se rend compte qu'il y a des variations naturelles autour d'une tendance mais qu'il y a aussi des variations exceptionnelles expliquées par des événements exceptionnels comme les pandémies, les guerres et les catastrophes naturelles. Il apparaît important de prendre en compte ce phénomène dans les projections démographiques à long terme. La modélisation classique est de modéliser la surmortalité due à ces événements sous la forme d'un produit d'une composante fréquence par une composante intensité. Certains experts de l'OMS estiment qu'il y a eu 31 événements de mortalité catastrophique sur les 420 dernières années soit en moyenne de 7.38 événements par siècle. Une modélisation de la surmortalité résultant des catastrophes naturelles est décrite dans un document de L. Devineau [10]. Le modèle peut s'écrire comme suit :

$$q(x, y, a, t) = q^{central}(x, y, a, t) + B_t(p)E_t(\lambda)$$

$$\text{logit}(q^{central}(x, y, a, t)) = \alpha_1(t) + a\alpha_2(t) + \sum_{i=1}^{M_x} \beta_i^l(t)^\top x_i + \sum_{j=1}^{M_y} \gamma_j^l(t)^\top y_j,$$

dans laquelle B_t est une loi de Bernoulli tirée à l'instant t dont la probabilité p correspond à la probabilité moyenne d'occurrence de la catastrophe. La loi E_t est une loi exponentielle dont le paramètre λ correspond à l'inverse de l'intensité moyenne de la catastrophe (la moyenne d'une loi exponentielle est l'inverse du paramètre).

L'ensemble des facteurs aléatoires décrit dans l'évolution du taux de mortalité sont représenté par un facteur Y_t^d .

Lien entre le taux de mortalité q et l'intensité de mortalité d

Plaçons dans le cas où Y_t^d est un facteur déterministe. La notion de taux de mortalité ici fait plutôt référence à une intensité de mortalité. Le taux de mortalité classique $q(x, a, Y_t^d)$ correspond à la probabilité d'un individu (x, a) de décéder au cours de l'année t et peut être relié à l'intensité de mortalité $d(x, a, Y_t^d)$ à l'aide de la fonction de survie.

En effet, la probabilité de survie $S_t(x, a, t + 1)$ d'un individu d'âge a et de trait x à la date t pendant une période d'un an s'exprime en fonction du taux de mortalité $d(a, x, Y_t^d)$ sous la forme suivante

$$S_t(x, a, t + 1) = \exp \left(- \int_t^{t+1} d(x, a - t + u, Y_u^d) du \right).$$

Or, l'expression de la survie d'un individu pendant l'année t en fonction du taux de mortalité q (au sens classique) est donnée par :

$$S_t(x, a, t + 1) = 1 - q(x, a, Y_t^d).$$

En identifiant cette expression avec l'expression précédente, on obtient :

$$\int_t^{t+1} d(x, a - t - u, Y_u^d) du = \ln \left(\frac{1}{1 - q(x, a, Y_t^d)} \right).$$

En faisant l'approximation d'une force de mortalité d constante sur une année, on obtient la relation suivante :

$$d(x, a, Y_t^d) \sim \ln \left(\frac{1}{1 - q(x, a, Y_t^d)} \right).$$

Remarque 7.3. Concernant la majoration de l'intensité par une intensité de mortalité \bar{d} , cela est possible si le taux de mortalité q est majoré par une valeur $\bar{q} < 1$. Cela est vérifié en pratique dans la mesure où l'on modélise la mortalité jusqu'à l'âge de 110 ans et au delà de 110 ans, le taux de mortalité est supposé constant de valeur strictement inférieure à 1.

Taux de mortalité en France en 2007

Le taux de mortalité en France peut être analysé sur les tables de mortalité nationale. On observe un taux de mortalité des femmes plus faible que celui des hommes. On représente ici la mortalité nationale des hommes et des femmes en fonction de leur âge.

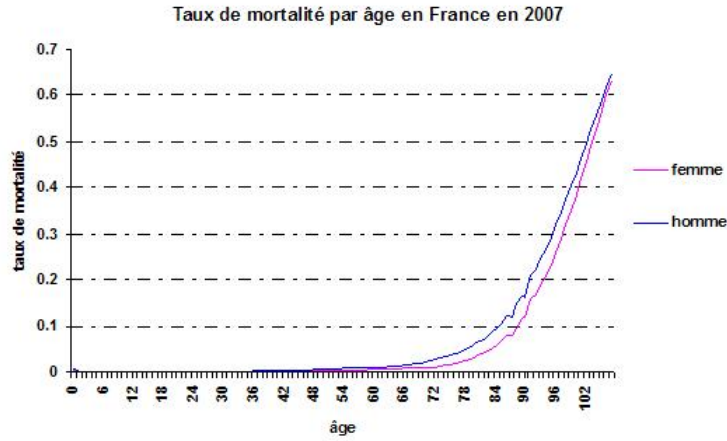


FIGURE 7.1 – Taux de mortalité par âge de la population française en 2007

On considère le modèle de Cairns à deux facteurs qui est un modèle de mortalité par âge. Ayant calibré le modèle, on déduit des scénarios d'évolution de ces deux courbes de mortalité dans le temps.

7.1.3 Taux de Natalité

Le taux de natalité correspond au rapport entre le nombre de naissances vivantes de l'année et la population totale moyenne de l'année (généralement la population en milieu d'année). L'intensité de natalité individuelle $b(x, a, Y_t^b)$ permet de quantifier la probabilité qu'une femme de trait (x, a) puisse avoir un enfant à la date t . En effet, comme pour la mortalité, on peut relier le taux de natalité $r(x, a, Y_t^b)$ à l'intensité de natalité $b(x, a, Y_t^b)$:

$$b(x, a, Y_t^b) \sim \ln \left(\frac{1}{1 - r(x, a, Y_t^b)} \right).$$

Par conséquent, l'analyse de l'évolution du taux de natalité permet de décrire l'évolution de l'intensité de natalité. Nous allons aussi modéliser la distribution des caractéristiques des enfants à leur naissance.

Étude de la natalité en France

L'indice synthétique de fécondité (ISF) correspondant au nombre d'enfants moyen d'une femme au cours de sa vie reflète assez bien le niveau de fécondité d'un pays. Cet indice synthétique de fécondité, de l'ordre de 2 en France, reflète un niveau assez élevé de fécondité par rapport aux autres pays européens. Il existe des projections de fécondité (méthode de Pearson) qui fournissent des hypothèses d'évolution de cet indice [14]. Il existe aussi des modèles statistiques à l'aide de processus autorégressifs qui décrivent l'évolution de l'ISF

[26]. Cependant, il faut caractériser le taux de natalité qui est plus difficile à modéliser.

Le taux de natalité dans les pays occidentaux a connu des variations très importantes durant le siècle passé. Du début du Baby Boom qui commence en 1945 jusqu'à la fin des "Trente glorieuses" vers 1975, le nombre de naissances à exploser alors que dès 1975, on observe une baisse spectaculaire de la natalité en Europe. En France, le taux de natalité, évalué par le nombre d'enfants pour 1000 femmes, était de 18.2 en 1964 et est descendu à son plus bas niveau en 1994 à 12.2. On observe en France une remontée du taux de natalité à partir de 1990 avec un taux qui atteint un niveau 12.9 en 2010 (Figure 7.2).

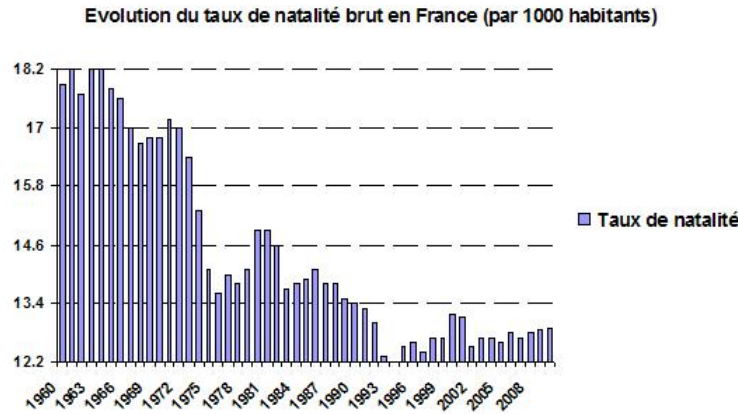


FIGURE 7.2 – Taux de natalité de la population française en 2007

Taux de natalité par âge

Il existe cependant des méthodes intéressantes d'ajustement analytique de courbe de natalité par âge [5, 13]. En effet, il est proposé une méthode paramétrique qui tient compte de l'indice synthétique de fécondité ISF , de l'âge moyen à la maternité \bar{a} et de la variance de l'âge à la maternité σ_a^2 . La plage de natalité est de $a = 15$ à $a = 50$ ans, et on modélise plutôt la courbe en $x = a - 15$. En effet, un modèle d'ajustement de courbe de natalité consiste à écrire que :

$$r(x) = \frac{ISF \cdot \lambda^p x^{p-1} e^{-\lambda x}}{\Gamma(p)},$$

où les paramètres sont donnés par :

$$\begin{aligned} x &= a - 15, & \bar{x} &= \bar{a} - 15, \\ \lambda &= \frac{\bar{x}}{\sigma_x^2}, & \bar{x} &= \frac{1}{ISF} \sum_{x=0}^{34} x f(x), \\ p &= \frac{\bar{x}}{\sigma_x^2}, & \sigma_x^2 &= \frac{1}{ISF} \sum_{x=0}^{34} (x - \bar{x})^2 f(x). \end{aligned}$$

Ainsi, en modélisant l'évolution dans le temps de l'indice synthétique de fécondité ISF_t , de l'âge moyen \bar{x}_t de maternité et de la variance $\sigma_{x,t}$, on pourrait en théorie avoir un modèle de projection de taux de natalité par âge $r(a, t)$ et avoir ainsi un modèle d'intensité de natalité $b(a, t)$. Cependant, ces méthodes d'ajustement ne sont pas très précises mais pourraient fournir des scénarios d'évolution de fécondité.

Déterminants de la natalité

Il est possible de déterminer des caractéristiques explicatives de fécondité. Généralement, on exprime le taux de natalité des femmes par âge mais il est possible de trouver d'autres caractéristiques explicatives. L'âge du mariage est un facteur déterminant de la fécondité. La manière de l'interpréter est que plus on se marie tard, plus on a de chances de divorcer et moins on a d'enfants.

Le niveau socioéconomique est aussi un facteur explicatif de fécondité. La fécondité n'a été que très peu prise en compte dans les travaux des économistes et on se rend compte que l'impact du niveau de revenu est important. G. Becker (prix nobel d'économie en 1992) modélise l'utilité des parents des pays occidentaux en utilisant le fait qu'ils sont soucieux d'avoir des enfants mais qu'ils sont aussi soucieux du bien être de leurs enfants [2, 3]. Dans les pays en voie de développement, les enfants sont considérés comme des placements rentables dans la mesure où il n'existe pas de système de retraites [24]. Les deux grandes conclusions de ces études économiques sont que :

1. Dans les pays occidentaux, lorsque les revenus augmentent, les ménages préfèrent avoir moins d'enfants mais consacrer davantage de dépenses à l'éducation de chaque enfant. G. Becker a démontré une relation inverse entre revenu familiale et le nombre d'enfants (en France, on observe une courbe en U).
2. Dans les pays pauvres, les pauvres cherchaient à avoir le plus d'enfants possibles puisqu'avoir des enfants est un gage de revenus dans le futur (absence de systèmes de retraites).

Ces conclusions permettent de comprendre le fait que la croissance du revenu dans les pays occidentaux s'est accompagnée d'un déclin de la fécondité dans les pays développés.

Il existe une théorie de la fécondité faisant intervenir le coût des enfants. En revanche, l'absence de données rend difficile d'analyser l'impact du revenu sur le taux de natalité. Cette modélisation est laissée à de futurs travaux de recherches.

Distribution des caractéristiques la descendance

Dans ce modèle, on prend en compte le fait que lors d'une naissance, les traits des enfants sont différents de ceux des parents. On décrit la méthode de prise en compte de cette mutation de trait avec un tirage des traits dans une distribution de descendance.

En ce qui concerne le trait lié au sexe, le descendant est un garçon avec une probabilité d'environ 0.512 ou une fille avec une probabilité d'environ 0.488. La mutation du sexe est donc facile à analyser. Le sexe de l'individu n'évoluera pas dans le temps. Le statut matrimonial est lui aussi immédiat dans la mesure où l'enfant naît célibataire. En revanche, le statut matrimonial pourra évoluer dans le temps avec un taux d'évolution de trait.

Afin de décrire l'évolution des traits de l'individu au cours de la vie, il faut considérer la structure des traits à l'instant initial. On peut faire le choix de fixer initialement certains traits à une valeur qui dépendra des caractéristiques des parents. En effet, il existe un déterminisme social et génétique dans la mesure où les caractéristiques des enfants dépendent souvent de celles des parents.

En ce qui concerne les traits reflétant le statut économique, le statut économique du père reflète assez souvent la catégorie sociale de l'enfant [11]. Une première idée consiste à fixer le trait de l'enfant égal au trait du parent (absence de mutation). Une deuxième méthode consiste à fixer le trait égal comme fonction du trait du père (mutation déterministe) ou de tirer le trait de l'enfant dans une distribution dont la moyenne correspond au trait du père (mutation éventuelle aléatoire à la naissance). On donne l'exemple de la CSP et du niveau de revenu qui illustrent les naissances sans mutations ou les naissances avec mutations aléatoires.

Concernant la CSP, on peut choisir de la fixer à la naissance comme la CSP du père qui reflète souvent le statut économique de l'enfant. On pourra tenir compte d'une éventuelle évolution de la CSP au cours de la vie de l'individu à l'aide du taux d'évolution de trait.

Pour le niveau de revenu, cela peut paraître moins raisonnable de le fixer égal au revenu du père et on peut considérer une éventuelle mutation à la naissance. Le revenu de l'enfant peut être tiré dans une loi normale centrée sur le niveau de revenu du parent, qui est tronquée sur un intervalle des revenus possibles $[0, \text{RevenuMax}]$. Il faut par ailleurs choisir la variance de cette loi normale qui pourra dépendre du niveau de revenu du parent : un parent

”riche” a de fortes chances d’avoir un enfant plutôt ”riche” et par conséquent la variance sera petite, tandis que pour un parent pauvre, la variance sera a priori plus élevée. Le niveau de revenu dont l’évolution est assez stationnaire sera lui aussi amené à évoluer durant la vie de l’individu. À titre d’exemple, on peut supposer une croissance de revenu avec un taux annuel dépendant par exemple de la CSP ou du niveau d’éducation de l’individu.

Par conséquent, à la naissance de l’individu, les traits seront déterminés (avec ou sans mutations) : certains traits n’évolueront pas (par exemple le sexe) et d’autres traits pourront évoluer dans le temps (par exemple la CSP, le niveau de revenu, le statut matrimonial...).

Ces études sont très complexes et ne seront pas abordées dans une première étude mais il est important de remarquer que le modèle permet de tenir compte de mutation de traits à la naissance et dans le temps.

Taux de fécondité par âge en France en 2007

Le taux de fécondité par âge correspond au nombre de naissance par femme ayant un âge donné. Ce taux permet de calculer la probabilité qu’une femme ayant un âge donné ait un enfant.

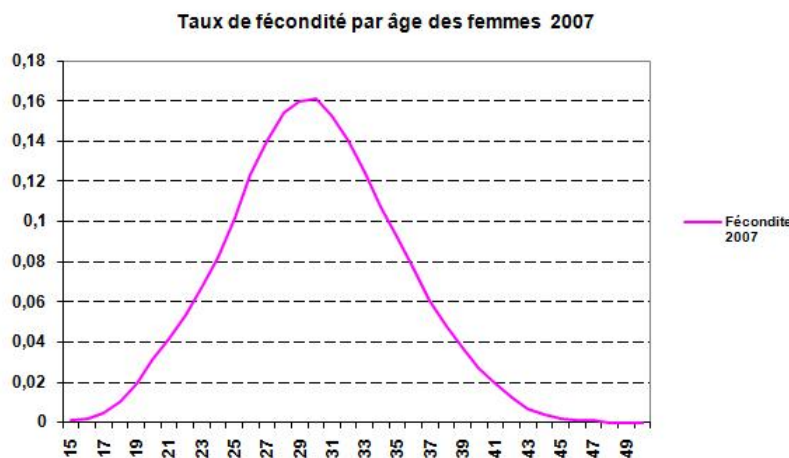


FIGURE 7.3 – Taux de fécondité par âge de la population française en 2007

On considère une natalité qui n’évolue pas dans le temps. L’absence de modèle pertinent pour la dynamique de courbe de natalité contraint à une modélisation statique. Cette extension sera l’objet de futurs travaux.

7.1.4 Taux d'évolution de traits

Les modèles classiques ne prennent généralement en compte que les événements de naissance et de décès. Cependant, les traits d'un individu sont amenés à évoluer dans le temps et il apparaît important de modéliser ce phénomène. Au cours de la vie d'un individu, divers événements peuvent apparaître comme un changement de statut matrimonial, l'obtention d'un diplôme, un changement de catégorie sociale. Ce taux d'évolution permet de modéliser l'occurrence de ces événements dans le temps. D'ailleurs, l'occurrence de ces événements peut dépendre des caractéristiques de l'individu.

Un premier exemple est le niveau de revenu dont on sait que l'évolution est assez stationnaire en France. La croissance des revenus se fait à un taux quasi constant depuis une trentaine d'année avec un taux annuel de 1.8% sur presque l'ensemble de la distribution de revenus. Il faut souligner la spécificité des revenus les plus faibles et des revenus les plus élevés qui ont un taux de croissance plus élevé.

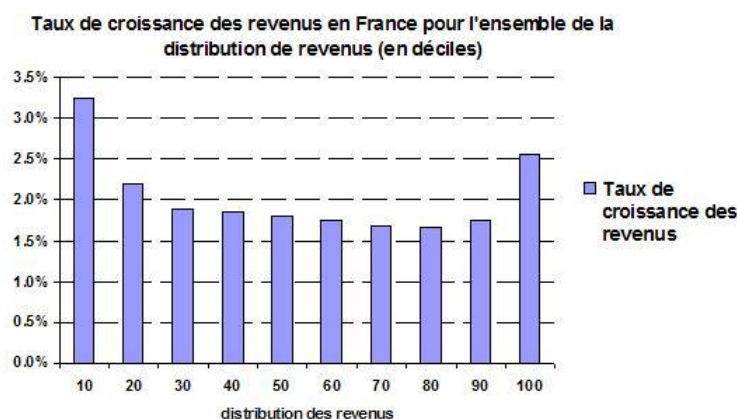


FIGURE 7.4 – Taux de croissance annuel des revenus en France pour chaque décile de la distribution des revenus

Des études plus complexes permettent de considérer la croissance du revenu individuel comme fonction de la catégorie sociale ou du niveau d'éducation. En ce qui concerne les inégalités de revenus, elles sont très stables depuis les cinquante dernières années. Pour information, l'indice de Gini est passé de 27.1% en 1995 à 27.2% en 2006.

Un deuxième exemple concerne le changement de statut matrimonial en France permettant d'illustrer la modélisation de l'évolution des traits. Des études démographiques décrivent les probabilités d'occurrence de divorce en fonction de certaines caractéristiques de l'individu. Il existe de nombreux déterminants du divorce comme l'âge au mariage, le niveau de

revenu de l'époux (épouse), le nombre d'enfants et le nombre d'année de mariage [21]. La modélisation du taux de divorce tiendra compte de ces résultats. On décrit sur la Table 7.1 l'évolution du taux de divorce en fonction du nombre d'année de mariage.

	1978	1988	1998	2008
Moins d'un an	0,1	0,1	0,1	0,6
1 an	3,0	4,5	3,4	6,3
2 ans	8,6	13,0	12,1	14,4
3 ans	11,9	17,1	18,4	21,7
4 ans	12,9	18,8	20,1	25,8
5 ans	13,3	17,3	22,0	27,3
6 ans	12,9	17,1	21,9	25,0
7 ans	12,4	16,4	21,1	22,6
8 ans	12,9	15,6	19,3	21,2
9 ans	11,0	14,5	18,0	19,1
10 ans	10,1	14,3	16,8	19,0
11 ans	9,2	13,4	16,4	17,3
12 ans	8,7	12,4	14,6	16,0
13 ans	8,1	11,9	14,0	15,6
14 ans	6,9	11,2	13,4	14,6
15 ans	6,5	10,4	12,1	13,9
16 ans	5,8	9,9	11,7	13,5
17 ans	5,4	9,9	11,2	12,9
18 ans	5,1	9,4	10,5	11,7
19 ans	4,3	8,5	10,5	11,8
20 ans	4,2	7,8	9,8	10,5
21 ans	3,7	7,2	8,7	10,1
22 ans	3,4	6,5	8,3	9,4
23 ans	3,1	6,1	8,2	9,2
24 ans	2,8	5,5	7,5	8,7
25 ans	2,4	4,7	7,0	7,9
26 ans	2,3	4,6	6,4	7,2
27 ans	1,9	3,6	5,6	6,7
28 ans	1,9	3,3	4,9	5,8
29 ans	1,4	2,8	4,4	5,2

TABLE 7.1 – Taux de divorce pour 1000 mariages en fonction de la durée du mariage

Ainsi, il est possible de tenir compte d'un taux d'évolution de statut matrimonial avec le passage du mariage au divorce. Avec le même genre de données, on obtient un taux de mariage, un taux remariage et un taux de veuvage. Le taux de veuvage dépend fortement des caractéristiques du conjoint. En pratique, on modélise le conjoint avec ses caractéristiques propres, on le fait évoluer simultanément et le veuvage intervient en cas de décès du conjoint. Ces taux d'évolution permettent de modifier le statut matrimonial au cours de la vie d'un individu.

7.1.5 Migration et immigration

Les flux migratoires jouent un rôle très important dans les projections démographiques. Nous décrivons les méthodes qui permettent de prendre en compte ces flux de population dans la modélisation et on analyse l'immigration en France.

Prise en compte de l'immigration

Les migrations des individus d'un pays vers un autre pays sont modélisées au niveau microscopique au même titre qu'un taux démographique. Concernant l'immigration, il existe les trois méthodes permettant de gérer les flux dans le modèle microscopique de dynamique de population (voir Chapitre 6).

La première possibilité est de considérer un taux d'immigration au même titre qu'un taux démographique individuel ce qui fournit la probabilité qu'un individu de caractéristiques données émigre ou bien qu'un individu immigré avec ces mêmes caractéristiques. Dans ce cas, l'immigration dépend des caractéristiques des individus de la population.

La deuxième possibilité est de modéliser le taux d'immigration comme étant un taux démographique global avec un nombre moyen de migrants. Il y aura des flux continus d'immigrants dans le temps et la distribution de l'âge et des traits des immigrants sera tirée suivant une loi fixée à l'avance.

La troisième méthode consiste à penser l'immigration comme une stratégie gouvernementale. L'idée est de laisser diffuser notre modèle de population et de l'arrêter à certaines dates (T_i) (pouvant correspondre à chaque année). À ces dates, on calcule le nombre d'immigrants nécessaires afin de satisfaire un critère que l'on s'est fixé au départ. L'âge et les traits des immigrants seront obtenus comme la réalisation d'une variable aléatoire dont la distribution sera fixée à l'avance. On met à jour la population en tenant compte du nombre d'immigrants nécessaires afin de satisfaire un critère donné. Ensuite, on relance le processus de diffusion avec la population modifiée par l'immigration comme condition initiale. Ces trois méthodes sont envisageables mais elles ne répondent pas aux mêmes besoins. Les deux premières méthodes considèrent un taux d'immigration dont l'évolution est fixée au début. La troisième méthode tient compte de l'évolution de la population dans le temps en adaptant la stratégie d'immigration de manière à satisfaire un critère. On utilisera les

deux premières méthodes pour les projections démographiques alors qu'on appliquera la troisième méthode pour l'étude du problème des retraites.

Taux d'immigration en France

On peut définir un taux d'immigration globale dans le modèle avec un flux continu dans le temps. L'idée est d'associer l'immigration à un processus de poisson dont le paramètre correspond au nombre moyen d'immigrés par an. On ajoute ce phénomène global au processus d'évolution qui met à jour la population dans le temps. L'âge et les traits des immigrants sont tirés dans une distribution définie au préalable.

On peut aussi introduire un taux d'immigration individuel à l'instar d'un taux démographique comme le taux de mortalité et le taux de natalité. Ainsi, en plus des deux événements possibles de naissance et de décès, il y a aussi un taux d'immigration. On considère $N(x, a, t)$ le nombre de personnes d'âge a et de trait x vivantes à la date t , et $D(x, a, t)$ le nombre de personnes d'âge a et de trait x qui ont décédé au cours de l'année t . On peut définir la probabilité qu'un individu d'âge a et de trait x à l'instant t immigrer en France par $im^+(x, a, t)$ et la probabilité qu'un individu d'âge a et de trait x émigre de France $im^-(x, a, t)$ de la manière suivante :

$$im(x, a, t) = \frac{N(x, a+1, t+1) - N(x, a, t) + D(x, a, t)}{N(x, a, t)}.$$

$$im^+(x, a, t) = \begin{cases} im(x, a, t) & \text{si } im(x, a, t) \geq 0 \\ 0 & \text{sinon} \end{cases}$$

$$im^-(x, a, t) = \begin{cases} -im(x, a, t) & \text{si } im(x, a, t) \leq 0 \\ 0 & \text{sinon} \end{cases}$$

On peut ainsi définir une intensité de migration $\rho^-(x, a, t)$ et une intensité d'immigration $\rho^+(x, a, t)$ qui sont données par :

$$\rho^+(x, a, t) = \log \left(\frac{1}{1 - im^+(x, a, t)} \right),$$

$$\rho^-(x, a, t) = \log \left(\frac{1}{1 - im^-(x, a, t)} \right).$$

L'objectif est d'adopter un modèle pour le taux d'immigration en analysant son historique au cours du vingtième siècle. La politique d'immigration a beaucoup varié au cours du 20^e siècle (voir Figure 7.5 et Figure 7.6) et il est quasiment impossible de considérer un modèle d'évolution de la force d'immigration consistant avec cet historique. En effet, l'Europe a été au centre de deux événements majeurs qui ont été les deux guerres mondiales. Les deux grandes guerres ont bouleversé les situations démographiques de l'ensemble des pays (avec

un taux de mortalité très élevé, des taux de naissances faibles et des migrations vers des pays plus stables). Ces périodes de guerre ont aussi eu des conséquences sur les années qui suivirent avec la période du baby boom pour la France qui dura une vingtaine d'années. L'immigration est un facteur qui est aussi affecté par certains événements majeurs qui ont touché d'autres pays que la France. Par exemple, vers les années 60/70, la France a été une terre d'accueil pour certains ressortissants de pays du Maghreb comme l'Algérie et la Tunisie qui ont connu leur indépendance pendant ces années. Certains démographes estiment que cette politique d'immigration a permis à la France d'avoir une situation démographique meilleure que celle de certains autres pays européens.

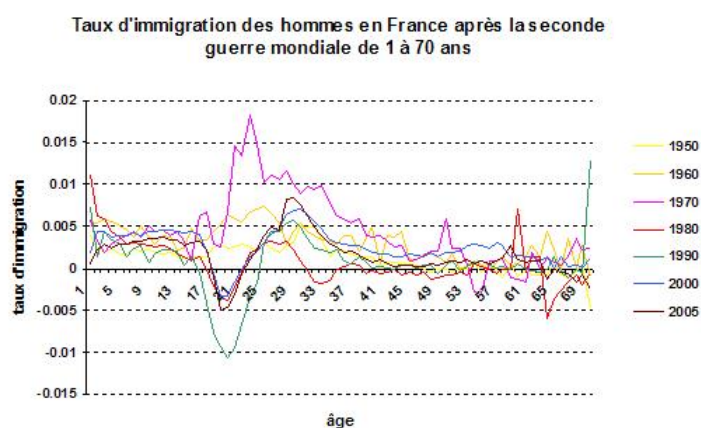


FIGURE 7.5 – Historique de la force d'immigration pour la deuxième moitié du 20^e siècle de 1 à 70 ans

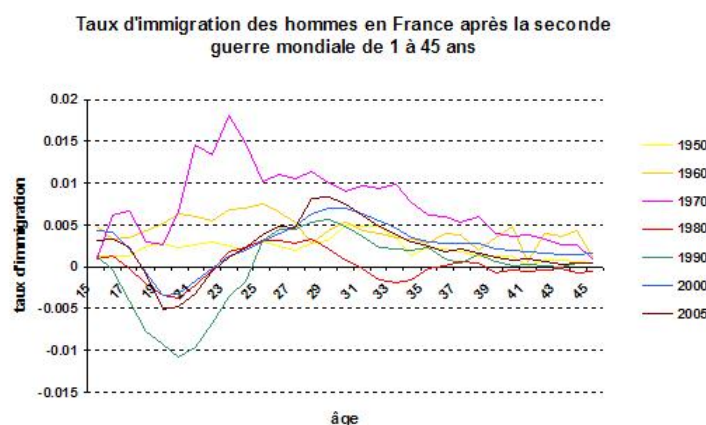


FIGURE 7.6 – Historique de la force d'immigration pour la deuxième moitié du 20^e siècle de 15 à 45 ans

Cependant, depuis les années 90, la politique d'immigration est restée quasiment la même avec des variations assez faibles (voir Figure 7.7 et Figure 7.8). On peut d'ailleurs noter un phénomène de migration des jeunes entre 19 et 26 ans qui vont s'installer dans d'autres pays comme aux États-Unis avec le "Brain Drain" (fuite des cerveaux). Ces migrations sont souvent des étudiants, des scientifiques ou des chercheurs qui s'installent à l'étranger pour avoir des meilleures conditions de vie, de travail ou de rémunération.

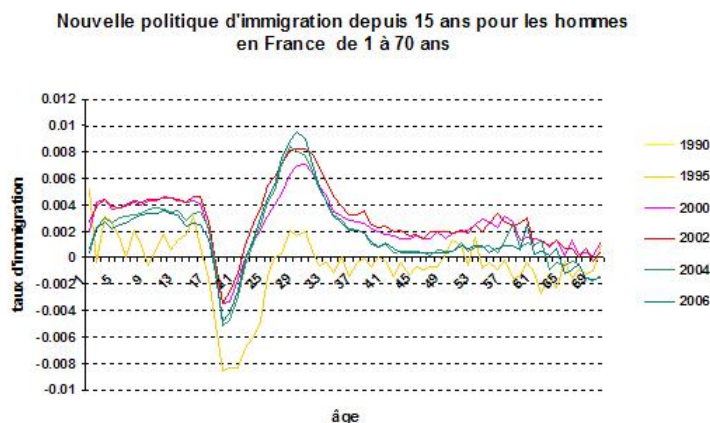


FIGURE 7.7 – Historique de la force d'immigration depuis l'an 2000 de 1 à 70 ans

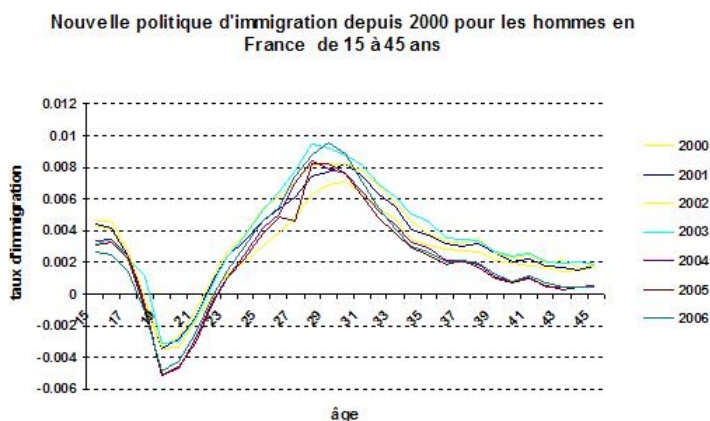


FIGURE 7.8 – Historique de la force d'immigration depuis l'an 2000 de 15 à 45 ans

Cela laisse penser que la régulation actuelle de l'immigration sera celle utilisée pour les prochaines années à venir. Ceci n'est bien sur que pure spéculation puisqu'un changement de gouvernement modifie généralement la politique d'immigration. Cette hypothèse permet de considérer le modèle d'immigration dans lequel le taux d'immigration est maintenu

constant pour les années à venir.

7.2 Projections démographiques de la population française

Dans cette partie, nous nous intéressons aux projections démographiques de la population française. Ayant analysé la situation démographique en 2007, nous simulons des scénarios d'évolution de la population dans le futur.

7.2.1 Bilan démographique de la population française en 2007

La situation démographique des pays occidentaux est assez alarmante avec un vieillissement global de la population résultant d'un taux de mortalité décroissant ainsi que d'un taux de fécondité assez faible. La France possède l'un des taux de fécondité les plus élevés d'Europe ce qui lui confère une situation démographique assez bonne. D'autres pays, comme l'Allemagne ou l'Italie, ont une situation démographique moins favorable avec un accroissement naturel négatif ce qui se traduit par une décroissance de la taille de la population.

À partir des données INSEE, on considère un échantillon de la population française en 2007 (constitué d'environ 62000 individus ce qui correspond à un facteur multiplicatif de $\frac{1}{1000}$) avec une bonne répartition des âges et des sexes par âges. Si on analyse la pyramide des âges en 2007 (voir Figure 7.9), on peut y noter les différents événements qui ont marqué le 20^e siècle. Le Baby Boom est visible sur la tranche d'âge 30/60 ans, la mortalité de la seconde guerre mondiale sur la tranche 60/75 ans. On remarque aussi le petit creux de la pyramide des âges vers les 90 ans causé par la première guerre mondiale (natalité plus faible) dont l'effet persévère 90 ans après.

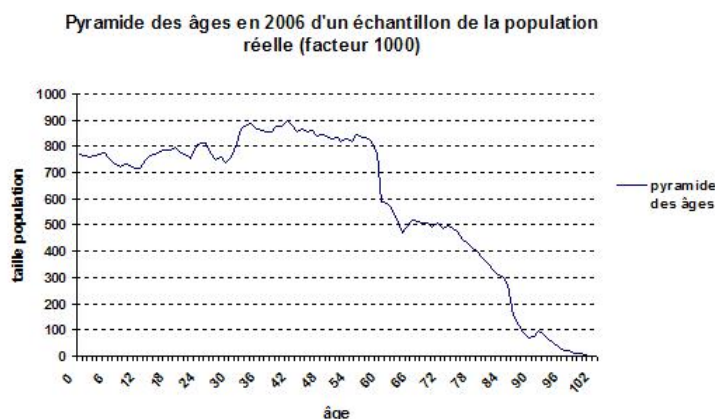


FIGURE 7.9 – Pyramide des âges de la population française (1/1000)

Considérant la situation démographique actuelle, l'INSEE et Alders [29] effectuent des prévisions de la situation démographique en 2050 que l'on décrit sur la Figure 7.10. Ces projections prévoient une augmentation de la proportion des personnes âgées avec l'entrée de la France dans l'ère du Papy Boom.

Comparaison des projections probabilistes de Alders et al. et de l'Insee

	Alders et al., 2007			Insee (Robert-Bobée, 2006)		
Hypothèses pour 2050	Limite basse (10 %)	Scénario central (médiane)	Limite haute (90 %)	Limite basse	Scénario central	Limite haute
Indice conjoncturel de fécondité (enfants par femme)	1,15	1,8	2,83	1,7	1,9	2,1
Espérance de vie des femmes, en années	85,5	89,7	94,1	86,5	89,0	91,5
Espérance de vie des hommes, en années	80,6	85,5	90,6	81,3	83,8	86,3
Migration nette (milliers)	-180	90	360	50	100	150
Population en 2050 (millions)	56,4	67,6	82,2	61,3	70,0	78,9
Proportion de 60 ans et plus (%)	26,6	32,4	38,0	28,0	31,9	36,3

FIGURE 7.10 – Projections démographiques INSEE et Alders pour 2047

La pyramide des âges va a priori s'inverser : les personnes issues de la génération du Baby Boom entrent dans le troisième âge et ces dernières sont susceptibles de vivre longtemps avec la forte baisse de mortalité.

Selon les hypothèses sur les taux démographiques, la population peut être amenée à croître ou à décroître. De plus, l'incertitude des prévisions est assez élevée puisque l'encadrement de la taille de la population donné par les instituts démographiques est assez large. En effet, à partir de modèles de taux démographiques, les démographes déduisent généralement un taux moyen ainsi que des limites basses et hautes pour chacun des taux démographiques. Ensuite, ils utilisent un modèle macroscopique déterministe qui prend en compte ces taux démographiques et déduisent des statistiques de projections démographiques de la population. On peut d'ailleurs remarquer les différences importantes qu'il existe entre les prévisions de chaque institut.

Dès lors, on utilise le modèle afin de décrire les prévisions démographiques de la population française dans le futur. On propose une analyse quantitative de l'évolution de la population française.

7.2.2 Étude qualitative de l'évolution de la population française

En biologie, dans le cadre d'une population asexuée, on cherche tout d'abord à savoir si elle est destinée à croître ou à s'éteindre. Pour cela, on suppose que les intensités d'évolution ne dépendent pas du temps et on définit un taux de reproduction net R_0 (voir Chapitre 6) donné par l'expression suivante :

$$R_0 = \int_0^{+\infty} b(a) e^{-\int_0^a d(\alpha) d\alpha} da.$$

On sait que pour une population structurée par âge avec une intensité de naissance $b(a)$ et une intensité de mortalité $d(a)$, le taux de reproduction net renseigne sur l'évolution "générationnelle" de la population dans le futur. En fait, ce taux correspond au nombre moyen d'enfants par individu au cours de sa vie dans l'hypothèse où les intensités restent constantes. Ainsi, la position de R_0 par rapport à 1 indique s'il y aura un renouvellement des générations ou non. Si $R_0 > 1$, il y a en moyenne plus d'un enfant au cours de la vie de chaque individu ce qui entraîne une explosion à long terme de la population. Si $R_0 < 1$, il y a extinction à long terme de la population car non renouvellement des générations. Enfin, si $R_0 = 1$, la population va atteindre un équilibre et on dit qu'on est dans le cas critique. On a l'intuition que la population française se trouve assez proche du cas critique sans explosion ni extinction.

Il est plus difficile de prendre en compte directement cette notion dans la mesure où la population est sexuée et que les données correspondent aux naissances par femme uniquement. En effet, le taux de reproduction net est un concept envisageable s'il existe un taux de naissance par âge et un taux de mortalité par âge identique pour l'ensemble des individus de la population. Dans cette modélisation, seules les femmes donnent naissance à des enfants alors que les hommes ne donnent pas "véritablement" de naissance, ce qui complique l'utilisation de ce concept. On ne peut pas modéliser une intensité de natalité $b(a)$ pour les femmes et une intensité de natalité nulle pour les hommes. On propose deux méthodes afin d'adapter cette notion à une population sexuée : la première est de considérer des taux démographiques individuels indépendant du sexe, et la deuxième est plus standard et consiste à ne considérer que le sexe féminin ainsi que les enfants de sexe féminin.

La première méthode est une étude qualitative qui n'a pas de justification mathématique rigoureuse mais qui reste cohérente avec la notion standard de taux de reproduction net. Si il y a autant de femmes que d'hommes dans la population étudiée (ce qui est à peu près le cas pour des populations humaines), le fait que la population soit sexuée avec une intensité de naissance par femme uniquement $b^F(a)$ pourrait revenir à considérer un taux de naissance par âge et par individu (quelque soit son sexe) égal à $b^*(a) = \frac{1}{2}b^F(a)$. Ce concept peut être justifié avec une approche d'un biologiste Charlesworth sur les populations sexuées [9].

Le ratio n'étant pas vraiment d'une femme pour un homme au sein de la population, et on aurait pu introduire une intensité de natalité égale à $b^*(a) = \frac{NbrFemmes(a)}{NbrHommes(a)+NbrFemmes(a)} b^F(a)$. Cependant, il faut aussi tenir compte des âges où un individu peut avoir un enfant et il y a une différence notable entre les hommes et les femmes. En effet, les femmes peuvent avoir des enfants jusque 45 ans (âge limité par la ménopause) alors que les hommes peuvent avoir des enfants bien plus tardivement. Ce phénomène ne peut pas être pris en compte du fait de l'absence de données de naissance par âge pour les hommes. Par conséquent, on fait le choix d'une intensité de naissance par âge pour une population asexuée donnée par $b^*(a) = \frac{1}{2}b^F(a)$. On définit aussi une intensité de mortalité $d(a)$ par individu indépendamment de son sexe. L'expression du taux de reproduction net dans ce cadre est donnée par :

$$R_0^1 = \int_0^{+\infty} b^*(a) e^{-\int_0^a d(\alpha) d\alpha} da.$$

La position du R_0^1 par rapport à 1, qui est sa valeur critique, va déterminer si il y a renouvellement des générations.

La deuxième méthode consiste à ne considérer que le sexe féminin avec son intensité de natalité et son intensité de mortalité. Le nouveau taux de reproduction R_0^2 correspond au nombre d'enfants par femme au cours de sa vie ce qui représente aussi l'indice synthétique de fécondité. Son expression est donnée par :

$$R_0^2 = \int_0^{+\infty} b^F(a) e^{-\int_0^a d^F(\alpha) d\alpha} da.$$

Puisque le sexe féminin est au centre de cette étude, le renouvellement des générations est assuré si en moyenne chaque femme donne naissance à au moins une fille. Il faut tenir compte que la distribution inégalitaire des sexes à la naissance. En moyenne, il naît 105 garçons pour 100 filles et il faudrait donc 205 enfants pour 100 femmes afin d'assurer le renouvellement des générations. Ainsi, la valeur critique de R_0^2 est de 2.05 correspondant au nombre d'enfants par femme nécessaire pour le renouvellement des générations.

On a calculé l'intensité de naissance $b(a)$ en considérant la probabilité $r(a)$ qu'une femme d'âge a donne naissance à un enfant (dont les données sont disponibles sur le site de l'INSEE). En effet, on utilise le lien taux/intensité :

$$b(a) = \log \left(\frac{1}{1 - r(a)} \right).$$

On fournit pour information les probabilités de fécondité, les intensités de naissance ainsi que les intensités de mortalité (pour un individu ou pour une femme) en 2007.

âge	Taux de fécondité $r(a)$	Intensité de fécondité $b(a)$	Intensité de mortalité $d(a)$	Intensité de mortalité $d^F(a)$
15	0.001	0.0010005	0.00019	0.00011
16	0.002	0.002002	0.00024	0.00018
17	0.005	0.005012	0.00035	0.00018
18	0.01	0.01005	0.00045	0.00023
19	0.019	0.01918	0.00044	0.00023
20	0.031	0.03149	0.00056	0.00026
21	0.041	0.04186	0.00055	0.00026
22	0.052	0.053408	0.00051	0.00017
23	0.065	0.06721	0.00058	0.00030
24	0.079	0.0823	0.00054	0.00027
25	0.097	0.10203	0.00060	0.00024
26	0.116	0.12329	0.00064	0.00031
27	0.131	0.14041	0.00063	0.00033
28	0.143	0.1543	0.00062	0.00033
29	0.148	0.16017	0.00062	0.00034
30	0.149	0.16134	0.00064	0.00035
31	0.142	0.1531	0.00064	0.00034
32	0.131	0.1404	0.00075	0.00036
33	0.117	0.1244	0.00075	0.00042
34	0.102	0.1076	0.00085	0.00051
35	0.088	0.0921	0.00091	0.00052
36	0.073	0.0758	0.00104	0.00058
37	0.059	0.0608	0.00114	0.00071
38	0.047	0.04814	0.00122	0.00076
39	0.0036	0.0366	0.00125	0.00086
40	0.027	0.02737	0.00142	0.00089
41	0.019	0.01918	0.00162	0.00103
42	0.012	0.01207	0.00179	0.00114
43	0.007	0.007025	0.00188	0.00123
44	0.004	0.004008	0.00214	0.00144
45	0.002	0.0020025	0.00248	0.00154
46	0.001	0.0010005	0.00262	0.00165
47	0.001	0.0010005	0.00289	0.00182

TABLE 7.2 – Probabilité de fécondité, intensité de natalité et de mortalité en 2007

Le calcul du taux de reproduction net selon les deux méthodes donnent $R_0^1 = 1.025$ et $R_0^2 = 2.042$ qui sont très proches de leur valeur critique. Ce taux de reproduction correspond au nombre d'enfants par individu (hommes et femmes) dans le premier cas et au nombre de filles par femme dans le deuxième cas. Cette quantité prend en compte le taux de natalité et le taux de mortalité jusqu'aux âges où la population cesse de se reproduire (pour les femmes, de l'ordre de 45 ans). Avec ces valeurs de taux de reproduction en 2007, si les taux démographiques étaient constants, la population devrait converger vers un équilibre. Ce raisonnement suppose normalement que les taux démographiques sont constants ce qui n'est pas le cas en réalité. Cependant, cette étude permet tout de même de souligner que la population française, comme la plupart des populations humaines, possède une structure naturelle ne permettant ni une explosion ni une extinction brutale mais plutôt une évolution vers un équilibre stable. L'étude qualitative ne permet pas de décrire la monotonie de la taille de la population dans le futur puisqu'on est très proche du cas critique. Il faut donc mettre en place une étude quantitative plus précise.

Remarque 7.4. *On peut facilement étendre la notion de taux de reproduction net au cadre microscopique comme cela été fait dans le chapitre 6 de la thèse. On définit un nouveau taux de reproduction net par âge et par traits donné par :*

$$\begin{aligned}\tilde{X}_t(x, a) &= (X_t^{x,a}, A_t^{x,a}), \\ R_{t_0}(X_{t_0}^{x,a}, A_{t_0}^{x,a}) &= \mathbb{E} \left(\int_{t_0}^{+\infty} b(X_t^{x,a}, A_t^{x,a}, Y_t^b) e^{-\int_{t_0}^t d(X_u^{x,a}, A_u^{x,a}, Y_u^d) du} dt \middle| \mathcal{F}_{t_0} \right),\end{aligned}$$

où le processus markovien $\tilde{X}_t^{x,a}$ décrit l'évolution dans le temps de l'âge et des traits d'un individu d'âge a et de trait x en date 0. Cette quantité permet de considérer le nombre d'enfants moyen d'un individu d'âge a et de trait x à la date t_0 jusque la fin de sa vie. Cette nouvelle définition prend en compte le fait que les facteurs démographiques évoluent dans le temps et que les individus sont aussi caractérisés par des traits qui peuvent évoluer.

7.2.3 Projections démographiques sans immigration

Après avoir analysé la situation démographique de la France en 2007, on génère des scénarios de la situation démographique dans le futur. On calibre les modèles associés à la mortalité des femmes et des hommes (mêmes sets de calibration que ceux du backtesting du chapitre 5) et on simule un scénario des tables de mortalité jusqu'à la date finale. On utilise les intensité de mortalité $d(a, t) = 1 - \exp(q(a, t))$ où q représente le taux de mortalité. Cela revient à supposer que la mortalité est uniforme sur une année donnée.

En l'absence de modélisation pertinente, on a maintenu le taux de natalité égal au taux de natalité en 2007 avec un indice synthétique de fécondité égal à 2.04 pour les simulations. Il

faudrait modéliser l'évolution de ce dernier pour des simulations encore plus intéressantes. Dans un premier temps, on ne tient pas compte des catastrophes naturelles ni des flux migratoires qui sont des phénomènes démographiques importants. Cette première étude permet de déterminer l'évolution naturelle de la population en l'absence d'apport migratoire. Considérant un échantillon fictif de la population française constitué d'environ 62000 personnes, on simule le processus d'évolution de la population avec ces taux démographiques (mortalité par sexe et natalité). L'idée est de voir l'évolution de la taille de la population dans le temps ainsi que la pyramide des âges à la date finale de projection.

Projections démographiques jusqu'en 2047

Les projections démographiques de la population française jusqu'en 2047 fournissent un panel de scénarios ainsi qu'un scénario moyen d'évolution. L'analyse des différents scénarios permet de décrire le comportement du modèle d'évolution.

Scénarios d'évolution de population Chaque scénario d'évolution de la population est obtenu par le tirage d'une trajectoire des intensités démographiques et par le tirage d'une trajectoire du processus d'évolution utilisant ces intensités démographiques. On représente ici le cas de la simulation de plusieurs scénarios d'évolution de la taille de la population jusqu'en 2047 (Voir Figure 7.11)

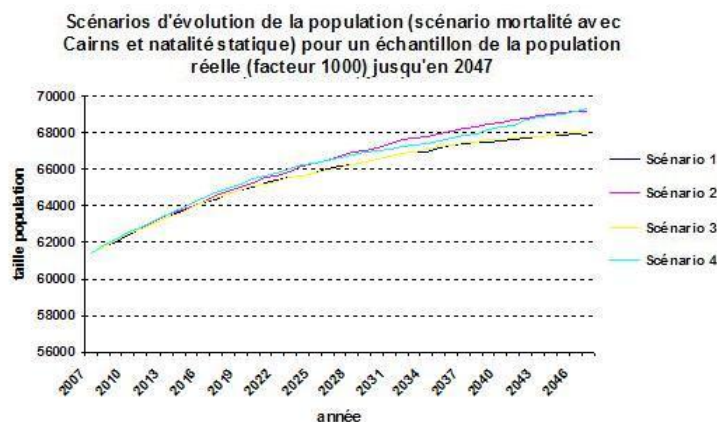


FIGURE 7.11 – Scénarios d'évolution de la taille de la population avec un scénario de mortalité jusqu'en 2047

On teste aussi l'évolution de la distribution des âges dans le temps. Pour cela, on représente plusieurs scénarios de pyramide des âges en 2027 et 2047 (voir Figure7.12 et Figure7.13)

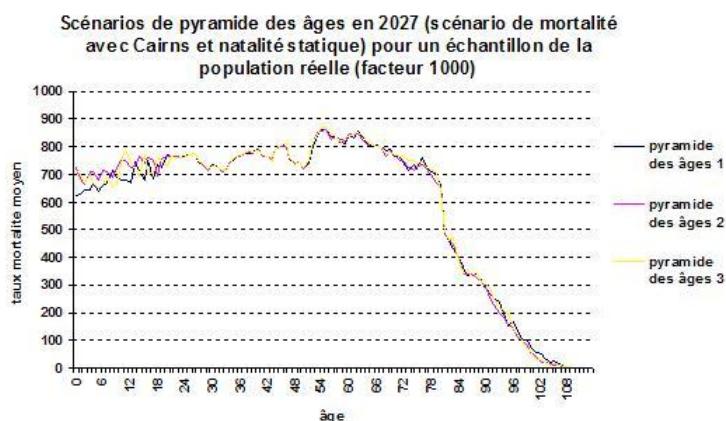


FIGURE 7.12 – Scénarios de pyramide des âges en 2027 avec un scénario de mortalité

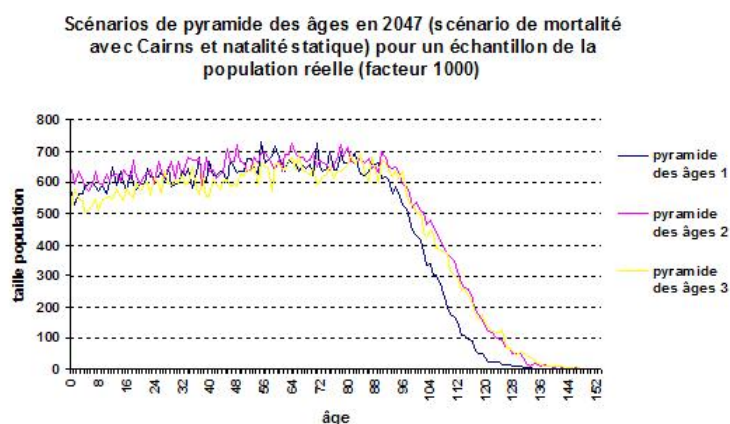


FIGURE 7.13 – Scénarios de pyramide des âges en 2047 avec un scénario de mortalité

En analysant ces graphiques, on se rend compte que notre modèle simule des scénarios de dynamique de population dont la variabilité augmente à mesure que le temps avance. Cette variabilité est due à la fois à l'aléa de la mortalité et à l'aléa du processus d'évolution de la population. La pyramide des âges en 2047 est elle aussi très variable d'un scénario à l'autre avec des écarts assez importants. De plus, en 2047, on se rend compte que la plupart des scénarios prévoient un âge maximum de plus de 130 ans ce qui est assez intéressant : si la mortalité continue à baisser de la sorte, l'âge maximum pourra repousser ses limites actuelles (124 ans avec Jeanne Calment en France). Ce modèle ne tient pas compte d'un éventuel âge limite que l'être humain ne pourra jamais dépasser. D'ailleurs, il est possible qu'en réalité la courbe des survivants se rectangularise au niveau de cet âge limite ce qui constitue un grand débat entre démographes, médecins et autres scientifiques s'intéressant

au sujet [22].

La structure de la pyramide des âges est très irrégulière alors qu'en pratique on souhaite avoir une pyramide des âges assez "smooth". Afin d'y parvenir, on peut supprimer la source d'aléa liée aux taux démographiques stochastiques en considérant une mortalité moyenne (moyenne sur 50000 simulations). Ainsi, il ne reste que l'aléa du processus de naissance et de mort ce qui permet de lisser les courbes. L'évolution de la taille de la population et les pyramides des âges sont plus régulières comme on peut le voir sur les figures Figure 7.14 à Figure 7.16.

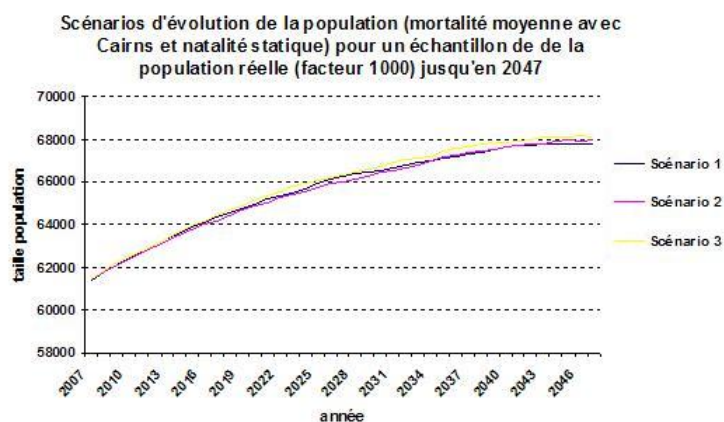


FIGURE 7.14 – Scénarios d'évolution de la taille de la population avec une mortalité moyenne jusqu'en 2047

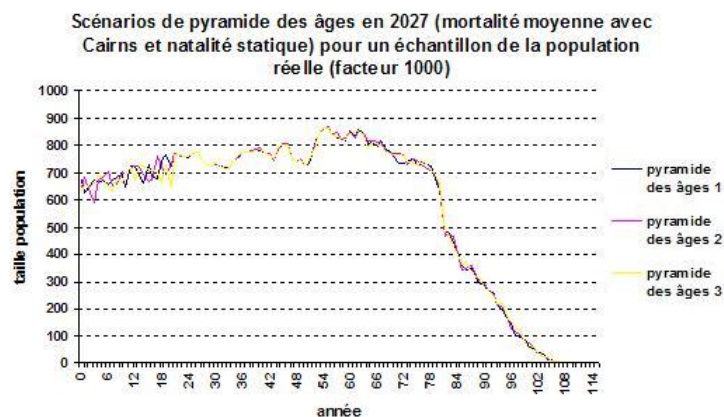


FIGURE 7.15 – Scénarios de pyramide des âges en 2027 avec une mortalité moyenne

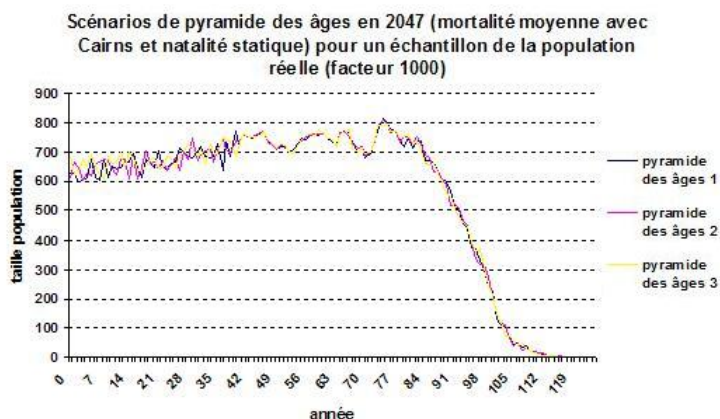


FIGURE 7.16 – Scénarios de pyramide des âges en 2047 avec une mortalité moyenne

Scénario moyen de projection L'aléa du processus d'évolution est bien moins important que celui des taux démographiques. En prenant une moyenne sur 50 scénarios de projections (tout en conservant une mortalité moyenne calculée avec 10000 scénarios), on peut encore régulariser nos projections en supprimant l'aléa du processus d'évolution. Cet aléa est assez faible ce qui traduit la stabilité du modèle de dynamique de population et il faut faire une moyenne sur assez peu de scénario pour converger vers un scénario moyen. L'évolution était déjà assez "smooth" mais la pyramide des âges se régularise énormément en faisant la moyenne sur les scénarios.

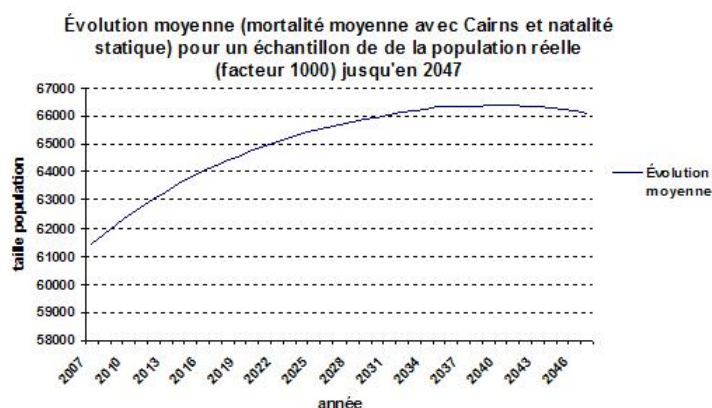


FIGURE 7.17 – Évolution moyenne de la taille de la population avec une mortalité moyenne jusqu'en 2047

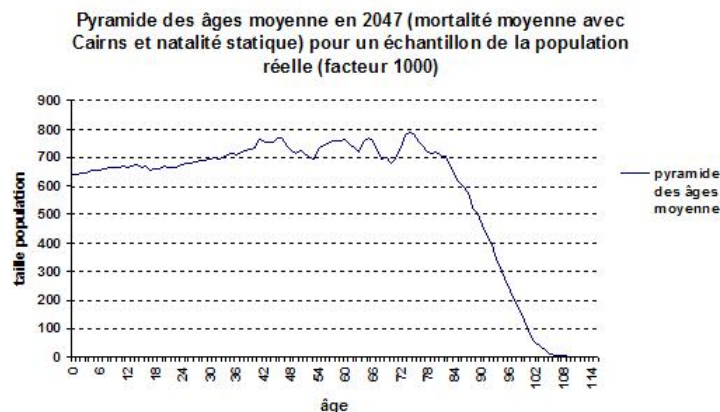


FIGURE 7.18 – Pyramide des âges moyenne en 2047 avec une mortalité moyenne

Le scénario moyen renseigne sur une prévision moyenne de l'évolution de la population qui devrait atteindre sa taille maximale vers 2042. Cependant, effectuer une moyenne de scénarios revient à ne considérer qu'une partie de l'information disponible. D'ailleurs, le scénario moyen ne reflète pas forcément un scénario d'évolution possible puisqu'il correspond à une agrégation de scénarios.

D'ailleurs, la pyramide des âges moyenne a un âge limite largement inférieur à celui d'un scénario d'une pyramide des âges. En effet, les âges élevés disparaissent en moyenne étant donné que le nombre de personnes d'un âge donné est faible et est spécifique à la réalisation du scénario. Lorsqu'on fait une moyenne sur un ensemble des scénarios, cela supprime ces cas exceptionnels de personnes vivantes à un âge très élevé. Le modèle prévoyait des scénarios pour lesquels certaines personnes dépasseraient des âges assez élevés (supérieur à 130 ans). Ainsi, afin d'analyser l'âge limite que le modèle fournit, il faudra tenir compte des scénarios différents scénarios plutôt que du scénario moyen qui ne fournit pas l'information voulue. La moyenne de l'âge maximal est de l'ordre 125 ans alors que l'âge maximal du scénario moyen est de l'ordre de 110 ans.

Projections démographiques jusqu'en 2097

Afin de décrire le comportement à très long terme du modèle, on s'intéresse à une évolution jusqu'en 2097. À l'horizon 2047, les scénarios d'évolution étaient tout de même assez proches les uns des autres pour des scénarios. Pour un horizon très lointain, on voit qu'on peut aboutir à des situations très différentes.

On considère à nouveau le cas d'une mortalité stochastique avec une natalité statique et aucune immigration. On analyse l'évolution de la population jusqu'en 2097 et on regarde

la pyramide des âges simulée en date finale (voir Figure 7.19 à Figure 7.22).

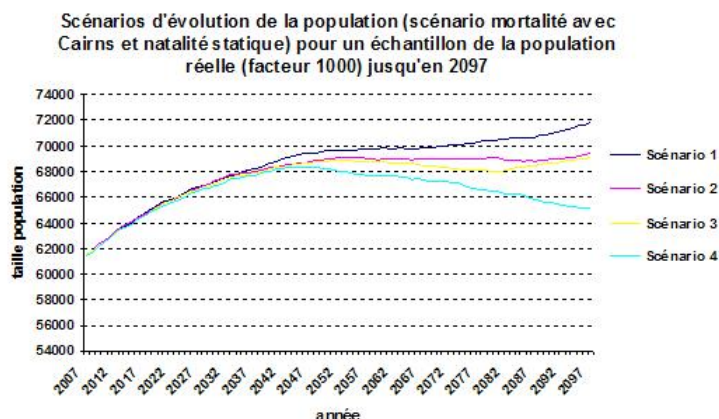


FIGURE 7.19 – Scénarios d'évolution de la taille de la population avec un scénario de mortalité jusqu'en 2097

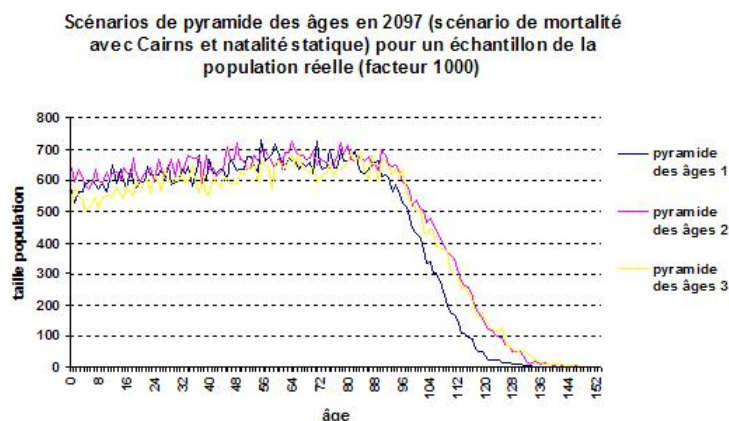


FIGURE 7.20 – Scénarios de pyramide des âges en 2097 avec des scénarios de mortalité

Les résultats apparaissent très intéressants du fait de la double stochasticité liée à l'aléa des taux démographiques et l'aléa du processus d'évolution. Pour des scénarios différents de mortalité, on obtient une évolution de population complètement différente : se trouvant proche du régime critique ($R_0^2 \sim 2.05$), la population peut continuer à augmenter, se stabiliser ou commencer à décroître. Il est important de comprendre que les facteurs R_0^1 et R_0^2 traduisent un comportement à long terme (à l'échelle des générations) mais qu'il est difficile de tirer de l'information sur l'évolution immédiate de la structure de la population. Par exemple, ce n'est pas parce que le facteur R_0^2 est inférieur à 2.05 que la population se

met immédiatement à décroître : il y a encore les personnes de plus de 47 ans qui vieillissent plus longtemps mais qui n'interviennent pas dans le calcul du taux de reproduction net.

On peut à nouveau prendre des taux démographiques moyens afin de régulariser ces courbes de projections.

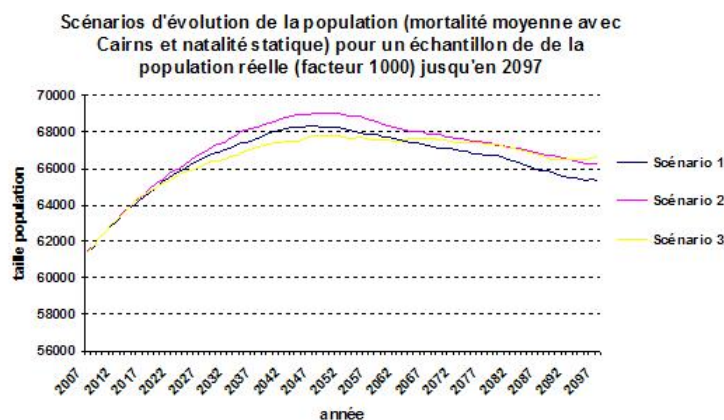


FIGURE 7.21 – Scénarios d'évolution de la taille de la population avec une mortalité moyenne jusqu'en 2097

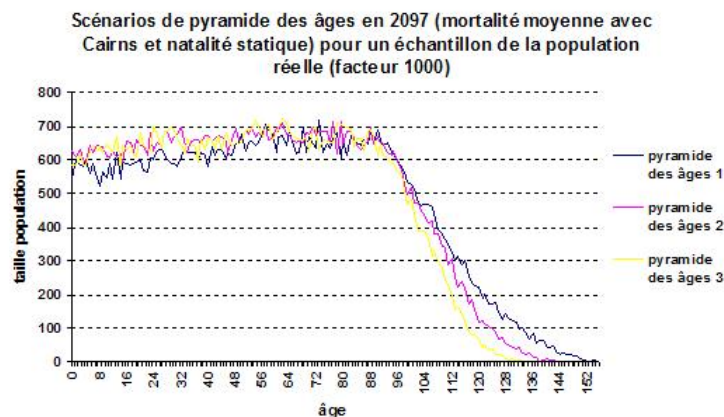


FIGURE 7.22 – Scénarios de pyramide des âges en 2097 avec une mortalité moyenne

Afin d'obtenir un scénario moyen d'évolution, on détermine une moyenne de scénarios de taux démographiques que l'on intègre dans le modèle. Bien que l'on considère des taux démographiques moyens, on observe différents types de scénarios puisque le processus d'évolution possède son propre aléa. Cependant, les comportements sont similaires pour tous les

scénarios avec une évolution croissante jusqu'en 2050 environ puis une décroissance de la taille de la population. Cela prouve que l'aléa des taux démographique était prépondérant et que celui du processus d'évolution est assez stable.

Ainsi, en générant un scénario moyen de mortalité sans immigration, la population française est amenée à décroître à partir de 2050 pour l'ensemble des scénarios du processus d'évolution. Cette évolution moyenne s'explique par le fait que les individus appartenant à la génération du Baby Boom vont décéder à mesure que le temps avance ce qui va impacter fortement la taille de la population.

Il apparaît très important d'analyser non seulement le scénario moyen mais aussi les scénarios extrêmes qui semblent donner des résultats assez différents. L'analyse de ces scénarios extrêmes sera très importante pour les gouvernements qui se doivent de prendre des mesures politiques concernant la démographie. On verra que le taux d'immigration est un enjeu majeur dans le vieillissement de la population et dans le paiement des retraites.

Analyse des projections démographiques

Après avoir fait une description qualitative des scénarios, on quantifie les projections démographiques afin de les analyser en détails.

Statistiques de projection

On introduit le concept du *potential support ratio* ou *dependency ratio* qui est une notion démographique permettant d'interpréter la proportion de la population adulte (15/64) ans par rapport aux personnes âgées (≥ 65 ans). Ce ratio sera très utile dans le système des retraites puisqu'il renseigne sur le nombre de personnes actives par personne inactive.

Statistiques en 2050	Limite basse 10%	Scénario médian	Limite haute 90%
Indice conjoncturel de fécondité	2.04	2.04	2.04
Taille de la population (millions)	67.79	69.05	70.59
Pourcentage des plus de 60 ans	37.4	39.3	40.7
Potential Support Ratio	1.45	1.53	1.64
Nombre total de morts (millions)	21.35	22.49	23.66
Nombre total de naissance (millions)	29.03	30.43	32.3

TABLE 7.3 – Statistiques de projections

On peut analyser les scénarios de projections au travers des statistiques de projections. En effet, on exploite l'information issue de l'ensemble des scénarios sous forme d'un tableau de statistiques. Ces statistiques fournissent une information centrale mais aussi des bornes supérieures et inférieures en terme de quantiles, avec les quantiles extrêmes à 10% et 90%. On tire des informations sur la structure démographique de la population dans le futur.

Dans le cas du scénario médian, on constate que la proportion des plus de 60 ans dans la population devient très importante avec environ 39.3% de la population. Le potential support ratio est de 1.53 actifs pour un inactif ce qui révèle un vieillissement majeur de la population. On peut aussi interpréter l'effet de "moyennisation" dans les statistiques de projections.

Statistiques en 2050	Limite basse 10%	Scénario médian	Limite haute 90%
Indice conjoncturel de fécondité	2.04	2.04	2.04
Taille de la population (millions)	67.76	68.09	68.4
Pourcentage des plus de 60 ans	39.7	39.9	40.1
Potential Support Ratio	1.49	1.50	1.51
Nombre total de morts (millions)	22.25	22.35	22.45
Nombre total de naissance (millions)	29.06	29.38	29.67

TABLE 7.4 – Statistiques de projections avec une mortalité moyenne

En considérant une mortalité moyenne issue du modèle et non plus des scénarios de mortalité, les simulations du processus de dynamique de population ont une variance assez faible ce qui confirme le fait que le modèle de dynamique est assez stable lorsque les taux démographiques sont fixés. En effet, si on analyse par exemple la taille de la population, on passe d'un intervalle de confiance de [67.79, 70.59] à un intervalle de confiance [67.76, 68.4] (voir Table 7.4). D'ailleurs, le fait que modifier la mortalité des individus va avoir un impact sur le nombre de naissances bien que le taux de naissance soit supposé constant.

Sans apport migratoire, les projections démographiques de la France en 2050 révèlent un vieillissement important de la population. Par rapport à la situation d'autres pays européens, la France se porte plutôt bien à condition que le taux de fécondité se maintienne au niveau actuel. Un taux de fécondité élevé permettrait de ralentir le vieillissement de la population.

Impact des catastrophes naturelles

Pour information, on effectue aussi une projection jusqu'en 2050 en tenant compte dans

notre modèle d'éventuelles catastrophes telle que les pandémies et les guerres (voir Table 7.5). Évidemment, ces scénarios ne font en aucun cas office de prévision mais plutôt de description d'éventuels scénarios avec une probabilité associée au modèle que l'on utilise.

Statistiques en 2050	Limite basse 10%	Scénario médian	Limite haute 90%
Indice conjoncturel de fécondité	2.04	2.04	2.04
Taille de la population (millions)	67.04	68.46	69.98
Pourcentage des plus de 60 ans	37.5	39.2	40.7
Potential Support Ratio	1.44	1.53	1.64
Nombre total de morts (millions)	21.68	22.97	24.18
Nombre total de naissance (millions)	28.89	30.34	32.01

TABLE 7.5 – Statistiques de projections avec d'éventuelles pandémies

L'analyse que l'on peut faire à la vue de ces statistiques est intuitive. En incorporant d'éventuelles catastrophes, il apparaît évident que la mortalité sera a priori plus élevée. L'ensemble des scénarios vont révéler une augmentation de la mortalité sans affecter pour autant la natalité. Par conséquent la taille de la population diminue en moyenne de plus de 500 000 personnes.

De plus, on a supposé que l'impact des pandémies était uniforme sur l'ensemble des âges ce qui ne modifie quasiment pas la proportion des personnes âgées dans la population. En revanche, la modélisation de ces catastrophes ainsi que la calibration du modèle sont à améliorer, mais ces résultats révèlent une influence non négligeable de la prise en compte de pandémies dans les projections à un horizon lointain.

On ne peut pas comparer ces résultats avec les prévisions INSEE étant donné qu'on n'a pas pris en compte les flux migratoires dans cette première modélisation. La politique d'immigration a elle aussi un impact important sur la situation démographique dans le futur comme on le verra dans la partie suivante.

7.2.4 Projections démographiques avec prise en compte de l'immigration

Dans cette partie, on intègre les flux migratoires à la modélisation. L'historique du taux d'immigration rend difficile de modéliser une évolution pertinente et on suppose dans un premier temps un taux d'immigration constant dans le temps égal au taux d'immigration en 2007.

Simulation avec une immigration statique

En considérant un taux d'immigration par âge statique égal à sa valeur en 2007, on peut estimer quelle sera l'évolution de la population ainsi que la pyramide des âges dans les années à venir (voir Figure 7.23 et Figure 7.24). Le concept d'immigration individuelle reste cohérent avec l'algorithme de simulation du processus d'évolution.

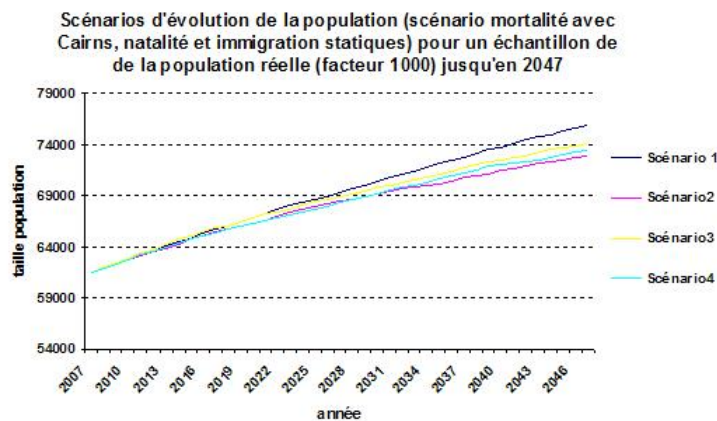


FIGURE 7.23 – Scénarios d'évolution de la taille de la population avec un scénario de mortalité jusqu'en 2047

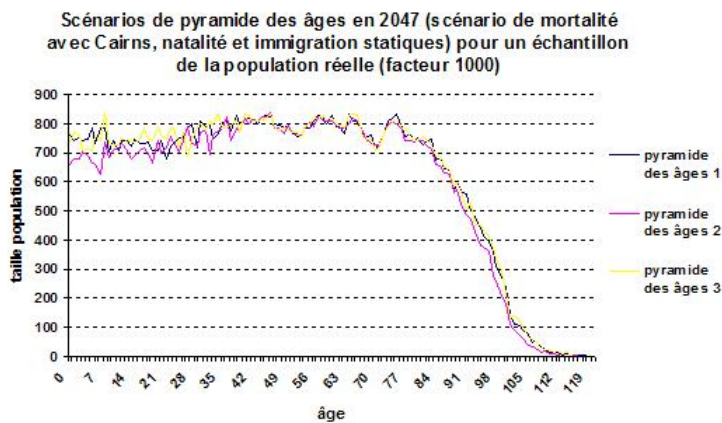


FIGURE 7.24 – Scénarios de pyramide des âges en 2047 avec différents scénarios de mortalité

Statistiques de projections et interprétations

On présente les statistiques de projections associées à cette évolution où l'on autorise une immigration statique, avec une éventuelle prise en compte de catastrophes naturelles (voir

Table 7.6 et Table 7.7)

Statistiques en 2050	Limite basse 10%	Scénario médian	Limite haute 90%
Indice conjoncturel de fécondité	2.04	2.04	2.04
Taille de la population (millions)	72.45	73.91	75.67
Pourcentage des plus de 60 ans	36	37.82	39.35
Potential Support Ratio	1.54	1.63	1.75
Nombre total de morts (millions)	21.5	22.68	23.91
Nombre total de naissance (millions)	31.07	32.47	34.41
Nombre d'immigrés net (millions)	2.816	2.97	3.146
Nombre d'immigrés par an (milliers)	65.49	69	73.17

TABLE 7.6 – Statistiques de projections avec une immigration statique

Statistiques en 2050	Limite basse 10%	Scénario médian	Limite haute 90%
Indice conjoncturel de fécondité	2.04	2.04	2.04
Taille de la population (millions)	71.56	73.12	74.94
Pourcentage des plus de 60 ans	36	37.8	39.2
Potential Support Ratio	1.54	1.63	1.75
Nombre total de morts (millions)	22.01	23.25	24.6
Nombre total de naissance (millions)	30.82	32.25	34.22
Nombre d'immigrés net (millions)	2.8	2.951	3.133
Nombre d'immigrés par an (milliers)	65.11	68.63	72.86

TABLE 7.7 – Statistiques de projections avec une immigration statique et d'éventuelles pandémies

Que l'on prenne en compte les éventuelles catastrophes ou non, les estimations révèlent une situation démographique complexe. Les immigrés s'ajoutent à la population et permettent d'augmenter le nombre de naissances. D'ailleurs, l'immigration modifie l'évolution long terme de la population dans la mesure où elle continue à croître dans les années 2050 alors qu'elle commençait à diminuer sans immigration.

Malgré cet accroissement de la population, on peut tout de même remarquer que la population continue de vieillir avec un pourcentage des plus de 60 ans qui reste élevé : 37.8% au

lieu de 39.3% dans le cadre d'une évolution sans apport migratoire. Par conséquent, cette politique d'immigration statique ne permet pas de changer significativement la structure démographique de la population. Cela s'explique du fait que la conjoncture démographique nous amène inéluctablement dans l'ère du Papy Boom.

Cependant, l'immigration peut être modélisée autrement qu'avec une immigration statique qui n'est pas réaliste sur un horizon lointain. En réalité, le taux d'immigration est relié à une politique d'un gouvernement qui n'est pas fixe dans le temps. Ainsi, il faudrait plutôt déterminer un taux d'immigration chaque année prenant en compte la structure de la population de la population et permettant de satisfaire un critère démographique.

7.3 Problème des retraites et stratégies possibles

La situation démographique des pays occidentaux est assez complexe dans la mesure où le vieillissement a un impact économique majeur sur la croissance et sur les problèmes des retraites. Le système de retraite français est intergénérationnel dans la mesure où la génération des actifs paient la retraite de la génération des inactifs. Il en résulte que le régime des retraites est fortement influencé par l'évolution de la situation démographique.

7.3.1 Introduction au problème des retraites

La France possède un système de retraite par répartition qui est basé sur un principe de solidarité intergénérationnelle. Avec les mutations du système de retraite d'un grand nombre de pays développés, on peut se poser la question de la viabilité d'un tel système. Avant la crise économique de 2008, un projet de changement du système de retraite avait été envisagé avec la création de fonds de pensions. En effet, la situation démographique actuelle laisse penser qu'un tel système ne pourra pas être efficace à long terme et que certaines mesures sont nécessaires.

L'État se propose de créer un fond spécial pour participer au financement des retraites futures. Afin de limiter le coût des retraites, l'État doit adopter une stratégie politique avec la mise en place de certaines réformes. Les solutions proposées actuellement sont la mise en place de fonds de pension, le recul de l'âge de départ de la retraite et l'adoption une politique d'immigration choisie.

Nous nous intéressons à l'aspect démographique du problème des retraites. Nous essayons d'étudier l'efficacité d'une immigration choisie ainsi que l'efficacité d'une réforme sur l'âge de départ à la retraite. Nous définissons le potential support ratio qui est une notion démographique intéressante renseignant sur la cotisation des retraites par la population active dans un système de solidarité intergénérationnelle.

Potential support ratio

Le "potential support ratio" ou "dependency ratio" est défini à l'échelle internationale comme le ratio du nombre de personnes d'âge compris entre 15 et 64 ans divisé par le nombre de personnes de plus de 65 ans. Actuellement, le niveau du potential support ratio est de l'ordre de 4 ce qui est satisfaisant pour le paiement des retraites. Cependant, les prévisions à l'horizon 2050 laissent supposer que le potential support ratio va descendre en dessous de 2 selon l'INSEE et vers 1.7 selon notre modèle. Les estimations de notre modèle sont légèrement plus pessimistes que celles de l'INSEE concernant le vieillissement de la population. Il existerait plusieurs moyens de maintenir une "situation démographique acceptable". Nous essayons de caractériser une situation démographique acceptable et de décrire les moyens possibles pour y parvenir.

Il existe de nombreux objectifs démographiques qu'un gouvernement souhaiterait atteindre. Un premier objectif pourrait être de garder une taille de la population croissante malgré une baisse de la natalité. Un autre objectif pourrait être de conserver la proportion des plus de 60 ans au dessus d'un niveau choisi. L'objectif le plus intéressant qui permettrait d'assurer le paiement des retraites jusqu'à un horizon donné serait de conserver un niveau de "potential support ratio" ou "ratio de dépendance" à un niveau donné. On va se focaliser sur ce dernier objectif relatif au paiement des retraites.

Le potential support ratio est une notion démographique qui n'est pas si représentative du point de vue économique. Cependant, des études révèlent qu'un potential support ratio de 3 serait satisfaisant dans la situation économique actuelle en considérant le taux de chômage, le produit national brut et d'autres facteurs économiques [29]. Les démographes et économistes estiment que le taux de chômage devrait baisser avec les départs en masse à la retraite des baby-boomer. Avec la crise économique de 2008, cette prédiction n'est plus tout à fait d'actualité puisque le taux de chômage n'a pas encore diminué malgré les premiers départs en masse à la retraite. Bien que la situation démographique idéale soit un potential support ratio égal à 4, on estime qu'un potential support ratio de 3 est une situation satisfaisante permettant d'avoir un système des retraites par solidarité intergénérationnelle viable.

7.3.2 Impact des flux migratoires

Les projections démographiques jusqu'en 2050 de la section précédente sont tout à fait alarmantes avec un vieillissement inéluctable de la population. Le pourcentage du nombre de personnes de plus de 60 ans ainsi que la statistique du niveau du potential support ratio (plus spécifique aux retraites) atteignent des niveaux critiques.

Dans un premier temps, on considère que le gouvernement ne dispose pas de richesse initiale (pas de financement de retraites) et que l'âge de départ à la retraite est fixé à 60 ans. On cherche donc à trouver une politique d'immigration afin de satisfaire un objectif à un horizon donné.

Remarque 7.5. *En fait, on peut faire un parallèle avec les actifs financiers dans la mesure où le taux d'immigration correspondrait à la stratégie que le gouvernement doit adopter afin de limiter les risques liés à l'évolution de la population. Le niveau d'immigration serait l'analogie du grecque delta en finance qui permet une couverture des risques liés à la variation du prix de l'actif.*

Un objectif potentiel est de maintenir un niveau de potential support ratio supérieur à 3 jusqu'à la date finale de projection. On cherche à caractériser la distribution optimale des immigrants de manière à satisfaire ce critère. En effet, l'âge des migrants n'est pas uniforme étant donné qu'il y a bien plus de migrants sur la tranche d'âge des 25/35 ans que pour les autres tranches d'âges (voir Figure 7.8).

La première distribution des migrants que l'on teste est la distribution des migrants en 2007 qui est celle que l'on a utilisé pour les projections démographiques (section 3.4). On peut voir que le taux d'immigration avec la distribution actuelle des immigrants par âge n'apparaît vraiment pas être un moyen envisageable ou même vraisemblable de satisfaire cet objectif. (voir Table 7.8).

Statistiques en 2050	Limite basse 10%	Scénario médian	Limite haute 90%
Indice conjoncturel de fécondité	2.04	2.04	2.04
Taille de la population (millions)	228	252	274
Pourcentage des plus de 60 ans	29.4	29.78	30.02
Potential Support Ratio	3	3	3
Nombre total de morts (millions)	26	27.3	28.6
Nombre total de naissance (millions)	54.6	56.8	59.1
Nombre d'immigrés net (millions)	141.2	162	179
Nombre d'immigrés par an (millions)	3.28	3.76	4.16

TABLE 7.8 – Statistiques de projections avec une immigration statique et d'éventuelles pandémies pour maintenir un potential support ratio de 3 jusqu'en 2050

En effet, il faudrait un flux migratoire d'environ 3.7 millions par an, ce qui apparaît totalement impensable. Une immigration non réfléchie (uniforme dans la classe d'âge de la population active 15/64 ans) est aussi invraisemblable comme le souligne la Table 7.9.

Statistiques en 2050	Limite basse 10%	Scénario médian	Limite haute 90%
Indice conjoncturel de fécondité	2.04	2.04	2.04
Taille de la population (millions)	229	240	253
Pourcentage des plus de 60 ans	29.5	29.7	30.12
Potential Support Ratio	3	3	43
Nombre total de morts (millions)	26.9	28	29.2
Nombre total de naissance (millions)	53.2	55.2	57.7
Nombre d'immigrés net (millions)	130	152	165
Nombre d'immigrés par an (milliers)	3.02	3.53	3.84

TABLE 7.9 – Statistiques de projections avec immigration uniforme des 15/64 ans et d'éventuelles pandémies pour maintenir un potential support ratio de 3 jusqu'en 2050

Une politique d'immigration uniforme sur la tranche d'âge des 15/64 ans fournit un niveau d'immigration totalement invraisemblable. On va donc chercher à modifier la distribution des immigrés en considérant le concept d'immigration choisie. Une immigration contrôlée qui n'autorise que l'immigration de personnes d'âges assez faibles permet d'optimiser l'impact de l'immigration. En effet, un jeune immigrant qui s'intègre rapidement dans la société fera parti de la population active et financera les retraites des inactifs plus longtemps. De plus, la venue de jeunes immigrants augmentent aussi le nombre de naissances. Il n'est cependant pas évident que les immigrés s'insèrent facilement dans la société et trouvent un travail rapidement. Il est d'ailleurs possible que l'immigration ait un impact négatif et des études précises doivent être menées.

Remarque 7.6. *Des études révèlent que le taux de natalité et le taux de mortalité des immigrés sont supérieurs à ceux des natifs [6, 12]. Les caractéristiques des immigrés étant différentes de celles des natifs, l'immigration modifie la structure de la population et ralentit le vieillissement de la population. On ne prend pas en compte les taux démographiques des immigrés bien que cela soit tout à fait pertinent.*

L'idéal serait de trouver une distribution optimale vraisemblable d'immigration qui permettrait de ramener le nombre d'immigrés à un nombre assez cohérent avec les politiques actuelles d'immigration. La distribution optimale mathématiquement serait a priori de n'autoriser que l'immigration de personnes âgées de 15 ans afin qu'elles cotisent tout au long de leur vie. Cela n'est d'ailleurs pas tout à fait exact lorsque l'horizon de projection est fixé étant donné l'influence du taux de fécondité sur le nombre de naissance, qui dépend aussi de l'âge de l'individu. Cependant, cette politique d'immigration n'apparaît pas vraiment réalisable en pratique puisqu'elle est totalement incohérente avec l'immigration actuelle et qu'elle se révélerait être une politique tout à fait inacceptable. Cette observation suggère

une politique d'immigration choisie en autorisant l'immigration d'une tranche d'âge donnée.

On décrit les résultats que l'on obtiendrait en mettant en place une politique d'immigration "choisie" sur une tranche d'âge. Dans un premier test, on autorise une immigration uniforme sur la tranche d'âge des 15/20 ans ce qui est très restreint. On considère ensuite une immigration sur la tranche d'âge des 15/40 ans qui apparaît être une politique plus raisonnable.

Statistiques en 2050	Limite basse 10%	Scénario médian	Limite haute 90%
Indice conjoncturel de fécondité	2.04	2.04	2.04
Taille de la population (millions)	114	121	127
Pourcentage des plus de 60 ans	26.1	26.4	26.7
Potential Support Ratio	3	3	3
Nombre total de morts (millions)	22.5	23.7	25
Nombre total de naissance (millions)	44.8	46.5	48.1
Nombre d'immigrés net (millions)	31.8	37	41.1
Nombre d'immigrés par an (milliers)	723	841	934

TABLE 7.10 – Statistiques de projections avec immigration uniforme des 15/40 ans et d'éventuelles pandémies pour maintenir un potential support ratio de 3 jusqu'en 2050

Statistiques en 2050	Limite basse 10%	Scénario médian	Limite haute 90%
Indice conjoncturel de fécondité	2.04	2.04	2.04
Taille de la population (millions)	114	119	124
Pourcentage des plus de 60 ans	22.3	22.6	22.8
Potential Support Ratio	3	3	3
Nombre total de morts (millions)	22.3	23.4	24.7
Nombre total de naissance (millions)	50.5	52.3	54.3
Nombre d'immigrés net (millions)	26.2	29.4	32.1
Nombre d'immigrés par an (milliers)	595	668	730

TABLE 7.11 – Statistiques de projections avec immigration des 15/20 ans et d'éventuelles pandémies pour maintenir un potential support ratio de 3 jusqu'en 2050

On voit qu'une immigration choisie permet de diminuer énormément le nombre d'immigrés nécessaire par rapport à la politique d'immigration actuelle. La moralité d'une politique d'immigration choisie est source de nombreux débats, on étudie quelles seraient les conséquences d'une telle politique sur la structure démographique. Cependant, les chiffres d'immigration restent encore trop élevés ce qui rejette la solution basée uniquement sur l'immigration choisie. On va considérer une réforme de l'âge de départ à la retraite associée à une politique d'immigration.

7.3.3 Impact de l'âge de départ à la retraite

Actuellement, de nombreux pays occidentaux repoussent l'âge de départ à la retraite afin de contrecarrer les effets du vieillissement de la population. L'agence de notation Moody's estime que "partir plus tard à la retraite augmentera le revenu disponible des salariés, réduira donc leur épargne ce qui stimulera la consommation donc la croissance". La Grèce qui a connu une crise économique majeure au cours de l'année 2010 a repoussé l'âge de départ à la retraite de 58 ans à 67 ans. En France, deux seuils légaux s'imposent : l'âge minimum de départ à la retraite est à 60 ans et l'âge pour partir avec une retraite à taux plein est à 65 ans. Actuellement, des réformes sont mises en place afin de repousser les âges de départ à la retraite.

On rappelle que le potentiel support ratio est défini à l'échelle internationale comme le ratio du nombre de personnes d'âge compris entre 15 et 64 ans divisé par le nombre de personnes de plus de 65 ans. On modélise l'impact d'une réforme sur l'âge de départ à la retraite sur la définition du potentiel support ratio. En effet, repousser l'âge de la retraite de x années permet aux personnes de cotiser x années de plus et de pouvoir partir à la retraite x plus tard. Par conséquent, si on considère une réforme qui repousse l'âge de la retraite de x années, il apparaît naturel de définir un nouveau potentiel support ratio correspondant au ratio des personnes d'âge compris entre 15 et $64 + x$ ans divisé par les personnes ayant plus de $65 + x$ ans.

Ainsi, si la France repousse l'âge de départ à la retraite de 5 années, on suppose que la définition du potentiel support ratio devient le rapport des 15 à 69 ans divisé par les plus de 70 ans. Cette nouvelle définition du niveau du potentiel support ratio permet de tenir compte des réformes sur les retraites en conservant la notion du nombre d'actifs par inactifs.

On modélise le report de l'âge de départ à la retraite de 5 ans avec la modification sous-jacente du potentiel support ratio. En considérant une immigration choisie, on cherche à voir quel va être le niveau annuel d'immigration suffisant permettant de satisfaire le critère démographique.

Statistiques en 2050	Limite basse 10%	Scénario médian	Limite haute 90%
Indice conjoncturel de fécondité	2.04	2.04	2.04
Taille de la population (millions)	84.7	90.1	96
Pourcentage des plus de 60 ans	30.3	31	31.6
Potential Support Ratio	3	3	3
Nombre total de morts (millions)	21.9	23.3	24.5
Nombre total de naissance (millions)	34.8	36	37.5
Nombre d’immigrés net (millions)	12.6	16.6	20.4
Nombre d’immigrés par an (milliers)	287	377	463

TABLE 7.12 – Statistiques de projections avec un report de 5 ans de l’âge de départ à la retraite, une immigration des 15/40 ans et d’éventuelles pandémies pour maintenir un potential support ratio de 3 jusqu’en 2050

Statistiques en 2050	Limite basse 10%	Scénario médian	Limite haute 90%
Indice conjoncturel de fécondité	2.04	2.04	2.04
Taille de la population (millions)	83.7	89.2	94.5
Pourcentage des plus de 60 ans	29.7	30.3	30.8
Potential Support Ratio	3	3	3
Nombre total de morts (millions)	21.9	23.1	24.3
Nombre total de naissance (millions)	35	36.3	37.7
Nombre d’immigrés net (millions)	10.9	14.9	18.5
Nombre d’immigrés par an (milliers)	247	338	420

TABLE 7.13 – Statistiques de projections avec un report de 5 ans de l’âge de départ à la retraite, une immigration uniforme des 15/20 ans et d’éventuelles pandémies pour maintenir un potential support ratio de 3 jusqu’en 2050

On observe que les niveaux d’immigration ont largement baissé pour atteindre un total de 15 millions sur 43 ans soit un nombre d’immigrés net égal d’un peu plus de 300000 par an. Ce niveau est à peine supérieur au niveau d’immigration des années 70 et est donc tout à fait envisageable. On en déduit qu’une réforme sur l’âge de départ à la retraite associée à une immigration choisie pourrait permettre de maintenir une situation démographique acceptable jusqu’en 2050 à condition de maintenir un taux de fécondité assez élevé.

7.3.4 Analyse des solutions proposées

Les tests que nous avons effectués précédemment servent à donner une intuition sur l'impact d'une stratégie ou d'une politique. En effet, nous avons vu que l'immigration ne permettait pas à elle seule de modifier la situation démographique de la France. La combinaison d'une immigration choisie et d'une réforme de l'âge de la retraite pourrait être une solution envisageable.

Nous avons analysé le problème des retraites d'un point de vue purement démographique ce qui n'est pas du tout suffisant. En effet, le problème est bien plus complexe et il faudrait prendre en compte l'impact économique du vieillissement de la population, la productivité du pays, le taux de chômage. De plus, il ne faut pas oublier que le gouvernement serait éventuellement prêt à financer les retraites alors qu'on a considéré un autofinancement du système intergénérationnel.

7.3.5 Formulation mathématique du problème des retraites

Nous avons essayé de décrire les solutions purement démographiques au problème des retraites en utilisant la notion de potential support ratio. Nous avons analysé l'impact du flux migratoire ainsi que de l'âge de départ à la retraite sur le niveau du potential support ratio.

Nous abordons l'aspect économique de ces solutions. L'idée est de définir économiquement ce qu'il faut entendre par une situation démographique "acceptable" d'un point de vue économique. Cela consisterait par exemple à donner un prix à une unité de potential support ratio et à voir de combien le gouvernement peut se permettre de voir le niveau actuel diminuer.

En effet, le gouvernement serait par exemple prêt à financer une partie des retraites avec une richesse initiale donnée (en faisant éventuellement un prêt) et il lui faudrait une stratégie politique qui lui permette de conserver un niveau de potential support ratio acceptable. Cette stratégie consistera a priori à fixer une richesse initiale, une politique d'immigration ainsi qu'un âge de départ à la retraite dans le but d'assurer le paiement des retraites jusqu'à une date donnée. D'un point de vue probabiliste, cela peut se voir comme un problème de contrôle stochastique sous contraintes. Nous proposons pour information une modélisation du problème d'optimisation sous-jacent.

On se place dans le cas d'une immigration globale avec un nombre moyen annuel d'immigrés. On peut poser $I_t = \int_0^t N(ds, da, dx)$ le processus ponctuel de Poisson décrivant le nombre d'immigrés jusqu'à la date t qui suit une loi de Poisson de paramètre ρt où ρ est le flux moyen annuel.

Le seuil toléré d'immigrés par an est noté A et le coût de la venue d'un immigré est noté c .

On note $\bar{Z}_t(dx, da) = Z_t(dx, da)1_{\{a > \beta\}}$ correspondant à la distribution des personnes ayant un âge supérieur à β et $\tilde{Z}_t(dx, da) = Z_t(dx, da)1_{\{15 < a < \beta\}}$ correspondant à la distribution des personnes ayant un âge compris entre 15 ans et β .

On peut définir l'âge de retraite maximum β_{max} avec par exemple $\beta_{max} = 70$ ans et le coût d'une unité de potential support ratio par p .

Ainsi, on cherche à regarder la fonction valeur J correspondant au coût d'une stratégie d'immigration associée à une réforme sur l'âge de départ à la retraite. On définit le processus d'immigration I et on note β l'âge de départ à la retraite. La fonction valeur est donnée par l'expression suivante :

$$J(\beta, I, T) = p \int_0^T \left(\frac{\langle \tilde{Z}_0, 1 \rangle}{\langle \bar{Z}_0, 1 \rangle} - \frac{\langle \tilde{Z}_s, 1 \rangle}{\langle \bar{Z}_s, 1 \rangle} \right) ds + cI_T.$$

On pourrait chercher à minimiser la moyenne de ces coûts sous certaines contraintes. L'idée est de se fixer au départ des valeurs acceptables pour le nombre maximal d'immigrés par an ainsi que l'âge maximal de départ à la retraite. Cela se ramène à l'étude de la fonction suivante :

$$v(A, \beta_{max}) = \min_{(\beta_t, I_t)} \mathbb{E}[J(\beta_t, I_t, T)], \quad \text{sous la contrainte} \quad I_t \leq A, \quad \beta \leq \beta_{max}.$$

On a décrit ce problème de contrôle stochastique pour information afin de voir qu'il est possible de faire une modélisation concrète du problème des retraites et de l'immigration.

7.4 Conclusion

L'évolution de la structure de la population des pays occidentaux est essentielle dans les études démographiques et économiques. Le modèle microscopique de dynamique de population permet d'obtenir un panel de scénarios possibles dont l'analyse fournit des statistiques sur les évolutions possibles. La propriété micro/macro du modèle permet d'avoir un comportement pertinent et cohérent à l'échelle macroscopique.

Dans cette étude démographique, nous avons considéré un modèle stochastique de mortalité par âge ainsi qu'un taux de natalité statique. Il existe bien évidemment de nombreuses extensions avec la prise en compte de taux démographiques stochastique qui dépendent à la fois de l'âge et de traits. Il faudrait aussi tenir compte de manière plus précise des

catastrophes naturelles dont la modélisation mérite de nombreuses extensions. Les modèles décrits dans les chapitres 5 et 6 permettent de traiter facilement un cadre général. Néanmoins, en l'absence de données pertinentes, nous avons dû effectuer les projections démographiques dans un cadre plus restreint.

Les projections démographiques révèlent que la population va vieillir inéluctablement et qu'il faut trouver une stratégie afin de limiter son impact. L'immigration est une des solutions considérées et il faudrait un nombre très élevé d'immigrés par an afin de modifier la structure démographique. Il faudrait mener des études précises sur l'immigration en analysant l'insertion des immigrés dans la société et en estimant le taux démographique spécifique aux immigrés.

Le système des retraites français est un système par solidarité intergénérationnelle qui dépend très fortement de la structure démographique du pays. Le débat actuel sur ce régime de retraite révèle qu'il est nécessaire de trouver des solutions rapides pour ne pas que le système implose. Les solutions apportées sont une politique d'immigration choisie et une réforme sur l'âge de départ à la retraite. Nous avons essayé d'analyser ces solutions d'un point de vue démographique en utilisant le concept du *potential support ratio*. Cette étude révèle que l'immigration choisie améliore significativement la situation démographique mais ne suffit pas à elle seule à résoudre le problème. Nous avons étudié l'effet conjoint d'une politique d'immigration choisie avec une réforme sur l'âge de départ à la retraite et il apparaît que cette solution pourrait permettre de sauver démographiquement le régime des retraites (du moins jusqu'en 2050).

Cette étude est bien évidemment naïve dans la mesure où le débat sur le système des retraites est bien plus complexe et dépend de la situation économique du pays avec le taux de chômage, la productivité du pays, le produit national brut (PNB) et autres facteurs déterminants. L'analyse économique concernant la viabilité du système des retraites est très intéressante et est laissée à des projets futurs de recherche. Il est d'ailleurs possible de formuler un problème mathématique de contrôle stochastique qui prendrait en compte le financement de l'État avec la prise en compte du coût économique de la venue d'un immigré et du vieillissement de la population.

Bibliographie

- [1] Antolin P. and H. Blommestein(2007), “Governments and the market for longevity-indexed bonds”, *OECD Journal : Financial Market Trends*, 4, 92.
- [2] Becker G. and H.G. Lewis (1973), “Governments and the market for longevity-indexed On the Interaction between the Quantity and Quality of Children”, *Journal of Political Economy*, 81, 279-288.
- [3] Becker G. (1993), “Human Capital : A Theoretical and Empirical Analysis, with Special Reference to Education”, *Third Edition, National Bureau of Economic Research*.
- [4] Bensusan H. and N. El Karoui (2009), “Microscopic Models for Populations Trait and Age Structures and Applications for Longevity problem”, Preprint.
- [5] Bourcier de Carbon P. (1989), “1959-1985 : la distorsion de la protection sociale au détriment des enfants”, *La jaune et la Rouge*, Octobre.
- [6] Brahim M. (1980), “La mortalité des étrangers en France”, *Population*, 35, 3, 603-622.
- [7] Bourgeois-Pichat J. (1951), “La mesure de la mortalité infantile. Principes et méthodes”, *Population*, 6, 2, 223-248.
- [8] Bourgeois-Pichat J. (1951), “La mesure de la mortalité infantile. Les causes de décès”, *Population*, 6, 3, 459-480.
- [9] Charlesworth N. (1994), “Evolution in Age-structured Population”, *Cambridge University Press*, 2 edition.
- [10] Devineau L. (2009), “La modélisation du risque de mortalité”, *Working Paper*, [http ://fr.milliman.com/publications/bulletin/index.php](http://fr.milliman.com/publications/bulletin/index.php).
- [11] Devaux M., F. Jusot, A. Trannoy and S. Tubeuf (2007), “Inégalités des chances en santé : influence de la profession et de l'état de santé des parents”, *Questions d'économie de la santé*, IRDES, 118, Février.
- [12] Dourgnon P., F. Jusot, C. Sermet and J. Silva (2007), “Etat de santé des populations immigrées en France : une approche multiniveaux”, *Working Paper, IRDES*.
- [13] Duchene J. and S. Gillet-de Stefano (1974), “Ajustement analytique des courbes de fécondité générale”, *Population et Famille*, Université Catholique de Louvain-Département de Démographie.

- [14] Ford D. and R.B.P Verma (1992), "Fertility Rates by Age of Mothers : Pearson Curve Type I vs Types II, III and Normal", *Presentation at the Federal Provincial Committee on Demography meeting*, Statistics Canada, December.
- [15] Fournier N. and S. Méléard (2004), "A microscopic probabilistic description of a locally regulated population and macroscopic approximations", *Annals of Applied Probability*, 14, 4, 1880-1919.
- [16] Imhoff E.V. and W. Post (1998), "Microsimulation methods for population projection", *Population*, 1, 97-136.
- [17] Girard A. and L. Roussel (1979), "Fécondité et conjoncture. Une enquête d'opinion sur la politique démographique", *Population*, 34, 3, 567-588.
- [18] Méléard S. and V.C. Tran (2008), "Adaptive Dynamics Models for Populations with Trait and Age Structures", *Journal of Mathematical Biology*, arXiv :0710.4997v1.
- [19] INED (1977), "La mortalité", *Population*, 32, 2, 284-338.
- [20] Jusot F. (2004), "Mortalité, revenu et inégalités de revenu en France : une analyse multi-niveaux", IRDES, Reference BDSP : 309602.
- [21] Lemennicier B. (1988), "Les déterminants du divorce", *Le Marché du mariage et de la famille*, Paris PUF, chapitre 3.
- [22] Levy M.L. (1996), "La rectangularisation de la courbe des survivants", *Colloque International de Sinaia*, 8, 576-579.
- [23] Monnier A. (2000), "La population de l'Europe : 1950-2050", *Populations et sociétés*, 353.
- [24] Razin A. and E. Sadka (1977), "Children as a capital good. A simple Model of Old Age Security with no Capital", *Population Economics*, 24.
- [25] Sauvy A. (1954), "Le vieillissement des populations et l'allongement de la vie", *Population*, 9, 4, 675-682.
- [26] Social Security Administration (2004), "A stochastic model of the long-range financial status of the OASDI program", Actuarial study 117, Septembre.
- [27] Schwarz K. (1968), "Influence de la natalité et de la mortalité sur la composition par âge de la population et sur l'évolution démographique", *Population*, 23, 1, 61-92.
- [28] Tran V.C. (2006), "Modèles particuliers stochastiques pour des problèmes d'évolution adaptative et pour l'approximation de solutions statistiques", Dissertation.
- [29] Toulemon L. (2007), "Projections démographiques pour la France et ses régions : vieillissement de la population et stabilisation de la population active", *Economie et statistique*, 408, 1, 81-94.
- [30] Von Foerster H. (1959), "Some remarks on changing populations", *The Kinetics of Cellular Proliferation*, In Grune & Stratton, editor, New York, 382-407.

- [31] Webb G.F., (1985), "Theory of Nonlinear Age-dependent population dynamics", *Questions d'économie de la santé*, New York : Marcel Dekker

BIBLIOGRAPHIE

Chapitre 8

Risques financiers des produits d'assurance vie

Introduction

Dans ce chapitre, nous étudions certaines applications financières de la modélisation des risques de taux d'intérêt, de mortalité et de longévité. Le domaine de l'assurance vie est un champ idéal d'applications aux modèles introduits dans les chapitres précédents de la thèse. En effet, les produits d'assurance-vie contiennent de nombreux risques dont le risque de longévité qui est lié à l'évolution de la longévité des individus ainsi que certains risques financiers [19]. Le risque de taux d'intérêt est aussi important dans la mesure où l'échéancier des flux s'étend sur un horizon lointain. La gestion des produits d'assurance vie est un enjeu majeur qui nécessite une compréhension et une modélisation de chacun de ces risques.

Motivations

Les assurances possèdent de nombreux produits de rentes ou de pensions qui contiennent un nombre important de risques. Le risque de longévité inhérent à ces produits provient de la difficulté à prédire l'évolution dans le temps du nombre d'assurés encore en vie.

Le risque de longévité à l'échelle nationale correspond au risque systémique de longévité qu'il est important de quantifier pour l'étude des produits d'assurance vie [4, 15]. Ce risque systémique est lié à l'incertitude de l'évolution dans le temps de la longévité des individus d'un pays. Le premier enjeu pour la gestion du risque systémique est l'estimation de la durée de vie d'un seul individu ce qui nécessite un modèle de mortalité. Un modèle probabiliste de mortalité permet de projeter des tables de mortalité nationale dans le futur en tenant compte de l'évolution de la mortalité dans le passé. Il faut modéliser l'évolution dans le temps de la mortalité nationale par âge et par sexe.

Le deuxième enjeu du risque systémique correspond à l'étude de l'évolution d'un échantillon de la population nationale dans le temps. Afin d'estimer la structure des flux du portefeuille de rentes, il faut estimer l'évolution du nombre de survivants du portefeuille dans le temps. Un modèle d'évolution de population permet de décrire des scénarios ainsi qu'un scénario moyen d'évolution de la population d'individus.

En plus du risque systémique de longévité, les compagnies d'assurance doivent aussi gérer le risque spécifique de longévité inhérent aux produits d'assurance vie [29]. En effet, les assurés n'ont pas les mêmes caractéristiques qu'un individu quelconque au sein de la population nationale et il ne faut donc pas appliquer les mêmes tables de mortalité. La mortalité des assurés est généralement plus faible que la mortalité nationale dans la mesure où leur niveau socioéconomique est souvent supérieur au niveau socioéconomique moyen. Ce risque spécifique de longévité est aussi appelé risque de base et il est assez difficile à estimer.

Le marché émergent des produits financiers visant à gérer le risque de longévité n'a pas encore trouvé sa place. De nombreux titres ou contrats (comme le "BNP Bond") n'ont pas eu le succès attendu puisqu'ils ne répondent pas aux attentes des différents intervenants et qu'ils ne permettent pas de couvrir le risque systémique et encore moins le risque spécifique d'un produit d'assurance vie (voir Chapitre 4). À cela s'ajoute la conjoncture actuelle avec un risque de crédit important puisque la contrepartie qui vend un contrat long terme peut faire défaut durant la durée de vie du produit.

Non seulement les compagnies d'assurance doivent gérer le risque de longévité, mais elles doivent aussi faire face aux risques financiers inhérents à leurs produits d'assurance vie dont elles ne sont pas forcément spécialistes. Le risque de taux d'intérêt est le risque financier principal sous-jacent aux portefeuilles de rentes mais certains produits comme les "variable annuities" contiennent de nombreux risques de marché [28]. L'évolution du niveau des taux d'intérêt est une source majeure de risque pour le bilan des assureurs puisque l'actif de ceux-ci est constitué en grande partie de produits de taux d'intérêt sur des périodes très longues (principalement en obligations) afin d'assurer le paiement des rentes. Les compagnies d'assurance cherchent un moyen de se débarrasser de ce risque de taux d'intérêt en le transférant à une autre institution qui accepterait de le supporter.

Présentation du chapitre

Dans une première partie, nous décrivons les modèles que nous utilisons pour estimer le risque de longévité en assurance vie. Nous nous intéressons plus précisément à la modélisation et à la gestion des risques de mortalité et de longévité d'un portefeuille de rentes ou de pensions dans la mesure où les assurances s'engagent à verser une rente aux assurés tout au long de leur vie. Ainsi, les modèles de mortalité et les modèles de dynamique de population

détaillés dans les chapitres 5 et 6 vont permettre de décrire les scénarios d'évolution des assurés du portefeuille. En supposant l'indépendance entre taux d'intérêt et longévité, il est possible de valoriser le portefeuille de rente comme la valeur actuelle nette du scénario moyen des flux de rentes.

De plus, les modèles de mortalité et de dynamique de population sont à l'échelle de l'individu et il est possible de tenir compte d'informations précises sur les individus dans l'estimation du risque de base du portefeuille. En étudiant les caractéristiques individuelles des assurés du portefeuille, nous avons estimé l'écart de mortalité de ces personnes avec le niveau national de mortalité qui représente un taux moyen. Cet écart de mortalité apporte de l'information sur le risque de base du portefeuille de rente. Malgré le concept de mutualisation des risques, les assureurs doivent gérer ce risque spécifique en considérant les caractéristiques particulières des assurés.

Nous nous intéressons aussi à la gestion des risques des produits de rentes et de pensions. Nous rappelons les produits qui ont été mis en place dans le but de gérer le risque de longévité et nous analysons l'ensemble des difficultés liées à la mise en place d'un marché de longévité pure. Les institutions possédant des produits d'assurance vie sont exposées naturellement au risque de longévité des assurés qui est couplé à un risque de taux d'intérêt sur un échéancier de durée très importante. Ces institutions cherchent à se débarrasser du risque financier tout en conservant le risque de longévité. Quant aux banques, elles sont spécialisées dans la gestion du risque de taux d'intérêt et pourraient supporter un risque de taux sur une maturité longue. En revanche, les banques ne veulent pas supporter un risque de longévité dans la mesure où elles ont une modélisation et une information limitées dans ce domaine. De plus, il n'existe actuellement pas de marché de longévité leur permettant de couvrir ce risque associant à la fois le risque systémique et le risque spécifique.

Nous nous rendons compte que de ce point de vue, les assurances et les banques ont potentiellement des intérêts en commun et qu'il serait possible de mettre en place un produit de transfert de risque qui répond aux attentes de ces deux entités. Le produit doit permettre de décorrélérer le risque de taux d'intérêt du risque de longévité et de transférer uniquement le risque financier. Nous rappelons dans un premier temps de nombreux concepts financiers classiques. Enfin, nous introduisons la *Longevity Nominal Chooser Swaption* (LNCS) qui est un produit spécifique que nous avons considéré comme le plus adapté à cette problématique. Ce produit permet à la banque et à l'assurance de gérer les risques dont elles sont spécialistes. Les assureurs modélisent de leur côté le risque de longévité de son portefeuille de rente et estiment des courbes de longévité. Les banques utilisent ces courbes de longévité ainsi qu'un modèle de taux d'intérêt pour évaluer le prix du produit. L'assurance pourra transférer la majeure partie de son risque de taux tout en conservant intégralement son risque de longévité. La banque quant à elle doit seulement gérer un risque de taux d'intérêt.

Pour l'étude de la *Longevity Nominal Chooser Swaption*, nous proposons d'utiliser les modèles de dynamique de population et de mortalité par âge et par traits décrits dans les chapitres précédents (Chapitre 5 et 6) afin d'analyser le risque de longévité du portefeuille de rentes. Du point de vue de la banque, elle pourrait utiliser le modèle de taux d'intérêt complexe (modèle de Wishart du chapitre 3) pour l'évaluation de ce produit¹. Cependant l'évaluation de ce produit exotique nécessite des méthodes numériques assez complexes à mettre en œuvre dans le modèle de Wishart. Par conséquent, nous avons choisi dans une première étude d'utiliser le modèle de Heath-Jarrow-Morton (HJM) [20] à deux facteurs qui est un modèle de taux d'intérêt assez simple permettant de prendre en compte certains facteurs de risques de ce produit (comme la corrélation des taux d'intérêt) et d'interpréter des résultats assez intuitifs sur ce produit.

8.1 Description de la modélisation du risque de longévité

Nous rappelons tout d'abord le modèle microscopique d'évolution d'une population détaillé au chapitre 6 de cette thèse qui s'inspire de travaux innovants en écologie [16, 17, 31, 32]. Dans le cadre de l'assurance-vie, nous ne nous intéressons qu'aux décès et aux éventuelles évolutions de traits dans le temps. En effet, la natalité ne présente pas d'enjeu véritable dans la modélisation des produits d'assurance vie. Nous décrivons ensuite le taux de mortalité individuel dont la présentation est faite au chapitre 5.

8.1.1 Processus d'évolution

Nous redéfinissons certains concepts liés à la modélisation microscopique de la dynamique de population. Notons $\mathcal{M}_P(E)$ l'ensemble des mesures ponctuelles sur E . Nous définissons un espace de traits χ muni d'une mesure de référence P et l'âge des individus est à valeurs dans \mathbb{R}_+ . Notons $\tilde{\chi} = \chi \times \mathbb{R}_+$ l'espace des caractéristiques des individus associant l'âge et les traits. Un individu sera donc caractérisé par une mesure de Dirac en $(x, a) \in \tilde{\chi}$.

Afin de décrire le processus d'évolution de la population, nous définissons sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ les éléments aléatoires indépendants suivants :

- La population initiale $Z_0 \in \mathcal{M}_P(\tilde{\chi})$ de taille finie est une mesure ponctuelle telle que $\mathbb{E}(\langle Z_0, 1 \rangle) < \infty$.
- $Q(ds, di, d\theta, dx')$ est une mesure ponctuelle canonique de Poisson sur $\mathbb{R}_+ \times \Xi := \mathbb{R}_+ \times \mathbb{N}^* \times \mathbb{R}_+ \times \chi$ et d'intensité $ds n(di) d\theta P(dx')$. ds représente le temps, $n(di)$ est la mesure de comptage sur \mathbb{N}^* qui sélectionne un individu, $d\theta$ détermine l'événement se produisant et $P(dx)$ est la mesure de référence sur χ .

1. Nous remercions l'équipe de recherche R&D du Crédit Agricole CIB pour nous avoir permis d'utiliser différentes bibliothèques utiles à l'évaluation du produit

- Les facteurs exogènes Y_t^d et Y_t^e sont des facteurs communs à l'ensemble de la population (à l'échelle de la population et non pas à l'échelle de l'individu), qui sont indépendants de Q , de Z_0 et qui sont indépendants entre eux.

Notons $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ la filtration canonique engendrée par Z_0 , Q et par les facteurs, Y^d et Y^e . Nous définissons un processus d'évolution de population Z à valeurs dans $\mathcal{M}_{\mathbb{P}}(\tilde{\chi})$ dont l'évolution est décrite par :

$$\begin{aligned} Z_t = & \sum_{i=0}^{N_0} \delta_{(X_i(Z_0), A_i(Z_0)+t)} + \int_0^t \int_{\Xi} 1_{i \leq N_{s-}} \left[-\delta_{(X_i(Z_{s-}), A_i(Z_{s-})+t-s)} 1_{0 \leq \theta < m_1(s, Z_{s-}, i, x')} \right. \\ & \left. + (\delta_{(x', A_i(Z_{s-})+t-s)} - \delta_{(X_i(Z_{s-}), A_i(Z_{s-})+t-s)}) 1_{m_1(s, Z_{s-}, i, x') \leq \theta < m_2(s, Z_{s-}, i, x')} \right] Q(ds, di, d\theta, dx'), \end{aligned}$$

où $X_i(Z_t)$ et $A_i(Z_t)$ correspondent respectivement aux traits et à l'âge de l'individu i de la population à la date t et où :

$$\begin{aligned} m_1(s, Z_{s-}, i, x') &= d(X_i(Z_{s-}), A_i(Z_{s-}), Y_s^d) k^d(X_i(Z_{s-}), A_i(Z_{s-}), x'). \\ m_2(s, Z_{s-}, i, x') &= m_1(s, Z_{s-}, i, x') + e(X_i(Z_{s-}), A_i(Z_{s-}), Y_s^e) k^e(X_i(Z_{s-}), A_i(Z_{s-}), x'). \end{aligned}$$

Le taux démographique $d(x, a, Y_t^b)$ représente le taux de mortalité qui dépend des caractéristiques de l'individu (son âge a et ses traits x) et d'un facteur Y_t^d commun à l'ensemble de la population. Ce facteur Y_t^d , éventuellement stochastique, caractérise l'évolution du taux de natalité dans le temps. De même, le taux d'évolution des traits est représenté par $e(x, a, Y_t^e)$.

Les événements affectant la population sont les décès et l'évolution de traits. Les occurrences sont modélisées à l'aide d'une mesure ponctuelle de Poisson. Entre deux événements, les individus vieillissent continûment avec le temps.

Remarque 8.1. *L'ensemble des propriétés mathématiques, l'ensemble des hypothèses sur les taux démographiques ainsi que la méthode de simulation du processus sont décrits au chapitre 6 de la thèse.*

8.1.2 Lien Micro/Macro

Contrairement à de nombreux modèles macroscopiques de dynamique de population, ce modèle est un modèle microscopique qui décrit l'évolution d'une population à l'échelle de l'individu. La modélisation microscopique présente de nombreux avantages même si elle est plus complexe à mettre en œuvre [22].

La modélisation microscopique permet de tenir compte d'informations précises sur chaque individu et fournit des informations plus riches en sortie. De plus, la modélisation est aléatoire et l'analyse des scénarios permet d'obtenir à chaque date un intervalle de confiance sur

le nombre de survivants. Contrairement à l'utilisation classique de modèles microscopiques en démographie, nous considérons la population complète des assurés dont nous connaissons précisément certaines informations et non pas un échantillon fictif de la population.

De plus, cette modélisation microscopique possède l'avantage majeur d'établir un lien micro/macro pertinent. Lorsque la taille de la population est grande, le modèle microscopique peut être approché par un modèle macroscopique pertinent. En effet, lorsque la taille de la population tend vers l'infini, le processus d'évolution converge vers un processus solution presque sûre d'une équation aux dérivées partielles stochastiques (voir Chapitre 6) qui généralise l'équation de McKendrick-Von Foerster classique en démographie [33, 34]. Cela prouve que le modèle microscopique possède de bonnes propriétés d'évolution.

L'intensité de mortalité d est modélisée afin de caractériser au mieux la mortalité des individus qui évolue dans le temps. En fait, ici d correspond plutôt à une intensité de mortalité. Le taux de mortalité q est décrit par un modèle de mortalité généralisant le modèle de Cairns par âge à deux facteurs à un modèle par âge et par traits. Nous rappelons le lien existant entre l'intensité de mortalité d et le taux de mortalité q .

8.1.3 Modélisation d'un taux de mortalité individuel

De nombreux modèles stochastiques de mortalité existent dont les plus classiques sont les modèles de Cairns [7, 8] et de Lee Carter [24, 25] avec l'extension à des modèles affines [30] éventuellement de type Wishart [18]. Nous étendons ce type de modèle à l'échelle de l'individu afin de tenir compte de toute l'information disponible et nous modélisons l'occurrence d'événements catastrophiques.

Certaines caractéristiques individuelles ont un impact important sur le niveau de mortalité comme la catégorie sociale, le niveau d'éducation, le statut matrimonial pour n'en citer que quelques unes [23]. La calibration du modèle sur les données de l'Échantillon Démographique Permanent (EDP) ainsi que l'impact de ces caractéristiques a été étudié en détails dans le chapitre 5. Rappelons la modélisation stochastique du taux de mortalité individuel.

Notons $\mathcal{A} = \{A_l, 1 \leq l \leq N\}$ la partition des classes d'âge où l'entier N correspond au nombre de tranches d'âge dans le modèle. Le taux de mortalité $q(x, y, a, t)$ correspond à la probabilité de décès à la date t d'un individu de traits $x = (x_i)_{1 \leq i \leq M_x}$ d'âge a vivant dans un environnement (ou une région) de caractéristiques $y = (y_i)_{1 \leq i \leq M_y}$. En fait, cette régression logistique sur chaque tranche d'âge peut s'écrire comme suit. $\forall l \in \{1, \dots, N\}, \forall a \in A_l$,

$$\text{logit}(q(x, y, a, t)) = \alpha_1(t) + a\alpha_2(t) + \sum_{i=1}^{M_x} \beta_i^l(t)^\top x_i + \sum_{j=1}^{M_y} \gamma_j^l(t)^\top y_j + B_t(p)E_t(\lambda).$$

La signification précise de chacune de ces variables et de ces paramètres, l'étude du modèle ainsi que sa calibration sont détaillées au chapitre 5 de la thèse. Rappelons tout de même l'interprétation de ces variables.

- Les variables y_i correspondent aux variables qui décrivent les caractéristiques d'environnement.
- Les variables x_i correspondent aux variables qui décrivent les caractéristiques individuelles explicatives (les traits) d'un individu i .
- Désignons $N_{x,i}$ le nombre de classes au sein de la variable individuelle x_i et $N_{y,i}$ le nombre de classes au sein de la variable globale y_i . Les variables peuvent être quantitatives avec une seule classe ou catégorielles avec un certain nombre de classes.
- Les coefficients $(\beta_i^l)_{\substack{1 \leq i \leq M_x \\ 1 \leq l \leq N}}$ et $(\gamma_i^l)_{\substack{1 \leq i \leq M_y \\ 1 \leq l \leq N}}$ correspondent aux coefficients de la régression logistique, les premiers étant associés aux traits individuels et les autres aux variables globales d'environnement.
- B_t est une loi de Bernoulli tirée à l'instant t dont la probabilité p correspond à la probabilité moyenne d'occurrence de la catastrophe. La loi E_t est une loi exponentielle dont le paramètre λ correspond à l'inverse de l'intensité moyenne de la catastrophe (la moyenne d'une loi exponentielle est l'inverse du paramètre).

L'événement des catastrophes naturelles affecte fortement la mortalité et il existe d'ailleurs des produits dérivés permettant un couvreur de ce type d'événements [10]. Ces phénomènes exceptionnels sont pris en compte en considérant une composante supplémentaire du taux de mortalité comme cela est fait dans les études biologiques ou certaines études démographiques. Par conséquent, il faut trouver des informations sur le taux d'occurrence de ces événements rares.

Les courbes de mortalité révèlent des variations naturelles autour d'une tendance ainsi que des variations exceptionnelles expliquées par des événements exceptionnels comme les pandémies, les guerres et les catastrophes naturelles. Il apparaît important de prendre ce phénomène en compte dans les projections démographiques à long terme. La modélisation classique consiste à modéliser la surmortalité due à ces événements sous la forme d'un produit d'une composante fréquence par une composante intensité. Certains experts de l'OMS estiment qu'il y a eu 31 événements de mortalité catastrophique sur les 420 dernières années soit en moyenne 7.38 événements par siècle. Ces données permettent de calibrer les paramètres p et λ décrivant l'occurrence et l'impact des événements catastrophiques.

Lien entre le taux de mortalité q et l'intensité de mortalité d

La notion de taux de mortalité est différente de l'intensité de mortalité du modèle de dynamique de population. Le taux de mortalité classique $q(x, a, Y_t^d)$ correspond à la probabilité d'un individu (x, a) de décéder au cours de l'année t et peut être relié à l'intensité de mortalité $d(x, a, Y_t^d)$ à l'aide de la fonction de survie.

En effet, la probabilité de survie $S_t(x, a, t + 1)$ d'un individu d'âge a et de trait x à la date t pendant une période d'un an s'exprime en fonction du taux de mortalité $d(a, x, Y_t^d)$ sous la forme suivante

$$S_t(x, a, t + 1) = \exp \left(- \int_t^{t+1} d(x, a - t + u, Y_u^d) du \right).$$

Or, l'expression de la survie d'un individu pendant l'année t en fonction du taux de mortalité q (au sens classique) est aussi donnée par :

$$S_t(x, a, t + 1) = 1 - q(x, a, Y_t^d).$$

En identifiant cette expression avec l'expression précédente, et en faisant l'approximation d'une force de mortalité d constante sur une année, on obtient la relation suivante :

$$d(x, a, Y_t^d) \sim \ln \left(\frac{1}{1 - q(x, a, Y_t^d)} \right).$$

8.2 Étude du risque de longévité d'un portefeuille de rente

De nombreuses études concernent le risque de longévité des portefeuilles de rentes en utilisant différents modèles de mortalité comme le modèle de Lee Carter [19], le modèle de Cairns [7] ainsi que des modèles affines [5, 13, 14, 27]. Nous associons le modèle de mortalité par âge et par traits au modèle microscopique d'évolution afin de tenir compte des caractéristiques individuelles. Nous décrivons la valorisation d'un portefeuille de rente et l'estimation de son hétérogénéité. L'analyse de l'évolution dans le temps du nombre d'assurés encore en vie permet d'évaluer le portefeuille de rentes.

Ces modèles à l'échelle de l'individu sont pertinents dans la mesure où ils tiennent compte d'informations importantes sur l'individu et que leur comportement est cohérent à grande échelle (voir Chapitres 6 et 7). Ces modèles sont adaptés à l'étude du risque de base d'un portefeuille puisqu'ils permettent de quantifier la mortalité spécifique des assurés et d'en tenir compte dans la valorisation du portefeuille. En effet, la mortalité des assurés étant généralement plus faible que la mortalité nationale, l'évolution du portefeuille sera elle aussi différente ce qui définit le risque de base.

8.2.1 Évaluation du risque systémique de longévité

On s'intéresse au risque systémique de longévité associé à un portefeuille de rentes en considérant en 2007 un portefeuille de 10000 personnes de 60 ans. En injectant les taux de mortalité prospectifs dans le modèle de dynamique de population, on obtient une évolution moyenne de la population. Cette évolution va correspondre financièrement à l'estimation de la série moyenne de flux qu'un assureur ou qu'un fond de pension s'engage à verser.

On représente pour une évolution du nombre de survivants, les quantiles à 5% et 95% de l'évolution de ce portefeuille qui sont calculés à partir de 10000 scénarios d'évolution du portefeuille. On distingue les hommes des femmes qui ont des niveaux de mortalité différents.

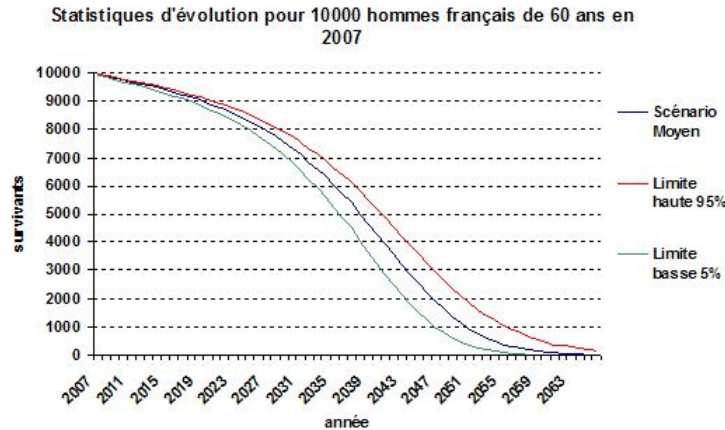


FIGURE 8.1 – Calcul des flux moyens à verser pour un portefeuille de rentes de 10000 hommes et estimation des flux extrêmes avec les quantiles à 5% et 95%

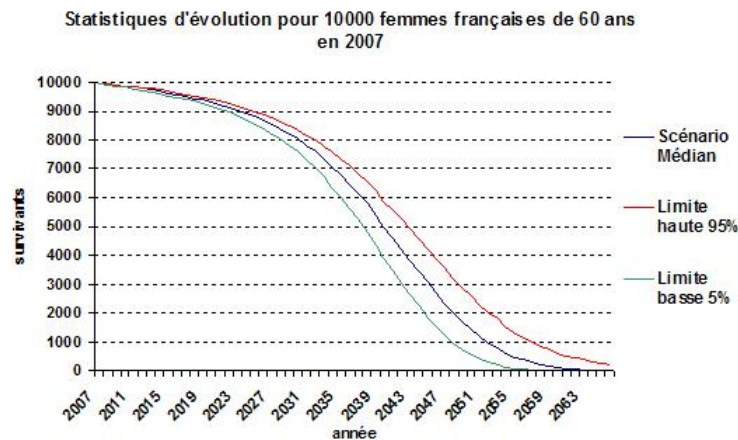


FIGURE 8.2 – Calcul des flux moyens à verser pour un portefeuille de rentes de 10000 femmes et estimation des flux extrêmes avec les quantiles à 5% et 95%

L'étude de ces graphiques renseigne sur le scénario central d'évolution mais aussi sur l'incertitude liée à l'évolution de la longévité dans les années à venir. En effet, les quantiles extrêmes d'évolution sont représentés par une courbe inférieure et une courbe supérieure qui forment un espace de confiance d'évolution. Pour des situations non extrêmes, la courbe

d'évolution du nombre d'assurés vivants appartient à l'enveloppe convexe délimitée par ces deux courbes.

On peut remarquer la différence qui existe entre la longévité des hommes et des femmes. Cet écart de longévité explique la différence de valorisation importante entre les rentes des hommes et celles des femmes. Par exemple, dans le scénario central, la population des hommes est diminuée de moitié en 2038 alors que celle des femmes est diminuée de moitié en 2040. On représente sur la Figure 8.3 les différences d'évolution entre les hommes et les femmes.

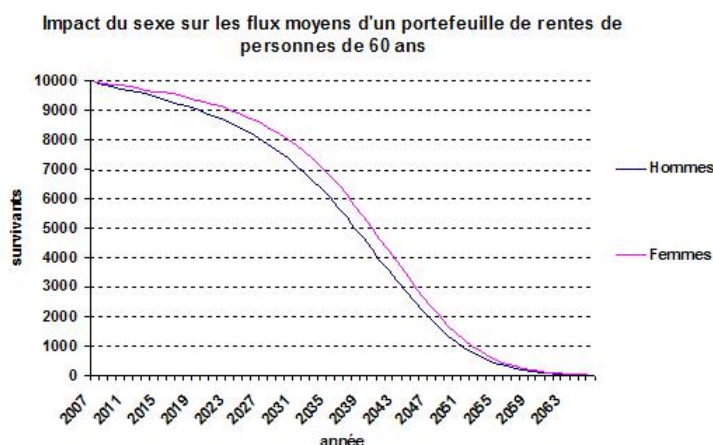


FIGURE 8.3 – Flux moyens d'un portefeuille pour 10000 femmes et pour 10000 hommes

8.2.2 Étude de l'hétérogénéité du portefeuille

Après avoir étudié le risque systémique du portefeuille de rentes, nous nous intéressons au risque spécifique (risque de base) résultant de l'hétérogénéité du portefeuille. Les caractéristiques des individus modifient la valorisation du portefeuille dans la mesure où le taux de mortalité aux âges élevés est largement impacté par certaines caractéristiques. Nous prenons l'exemple d'une population fermée de 10000 assurés de 60 ans qui ont tous une même caractéristique donnée comme cela pourrait être le cas pour un portefeuille de rentes. Nous analysons l'influence des caractéristiques sur la mortalité aux âges relativement élevés puisque les assurés ont plus de 60 ans. Nous comparons les séries de flux de rentes selon les valeurs prises par les caractéristiques en distinguant le cas des hommes de celui des femmes. Cette procédure permet d'estimer l'influence de la valeur du trait sur la valorisation de ce portefeuille de rentes.

Influence de la catégorie socioprofessionnelle (CSP)

Le niveau socioéconomique d'un individu peut être modélisé de plusieurs façons. En Angleterre, les compagnies d'assurance utilisent souvent le système "postcode" qui estime le niveau socioéconomique de l'individu en fonction de son lieu de résidence [3]. On considère ici la catégorie socioprofessionnelle (CSP) pour laquelle on a pu se procurer des données et on évalue l'impact de cette variable sur la valorisation des portefeuilles de rentes. On précise que l'on ne tient pas compte d'une éventuelle évolution de la catégorie sociale dans la mesure où elle est caractérisée à l'âge de la retraite par la profession exercée durant la vie de l'assuré.

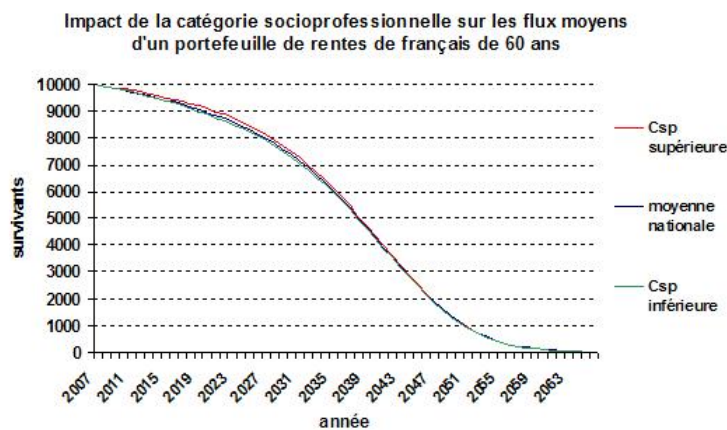


FIGURE 8.4 – Flux moyens d'un portefeuille constitué de 10000 hommes de 60 ans qui ont tous une CSP donné

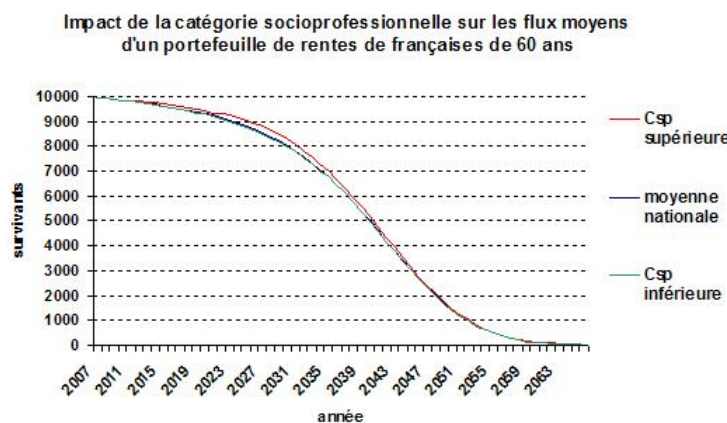


FIGURE 8.5 – Flux moyens d'un portefeuille constitué de 10000 femmes de 60 ans qui ont tous une CSP donné

On a opté pour une classification des catégories socioprofessionnelles prenant en compte l'indice standardisé de mortalité par CSP produit par l'INSEE

- inactifs (hors retraités) et non déclarés (données non exploitables) : CSP_0
- catégories "supérieures" (cadres et professions intellectuelles supérieures, artisans, commerçants, chefs d'entreprise, professions intermédiaires, agriculteurs) : CSP_1
- catégories "inférieures" (employés et ouvriers) : CSP_2

Ces graphiques révèlent que la catégorie socioprofessionnelle n'a pas une influence importante sur la valorisation du portefeuille de rentes que ce soit pour les hommes ou pour les femmes. On savait que la structure des catégories socioprofessionnelles avait peu d'influence sur les femmes (chapitre 5). On remarque ici que l'influence est aussi limitée pour l'homme car on se focalise sur les âges élevés. La catégorie socioprofessionnelle a une importance dans la longévité de l'homme surtout parce qu'elle diminue sa mortalité dans les âges actifs (jusque 65 ans). Après 65 ans, les écarts diminuent très rapidement pour ne plus être vraiment significatifs.

Influence du niveau de diplôme

L'impact du niveau d'éducation est mesuré par le niveau de diplôme. On distingue les individus ayant un niveau de diplôme supérieur à un BAC + 2 des individus ayant un diplôme inférieur à ce niveau. Le chapitre 5 révèle une influence du niveau d'éducation sur le niveau de mortalité, et ceci jusque des âges assez élevés. Ici encore, la variable d'éducation est une variable fixe dans la mesure où elle ne change pas au cours de la vie adulte d'un individu (à partir de 30 ans du moins).

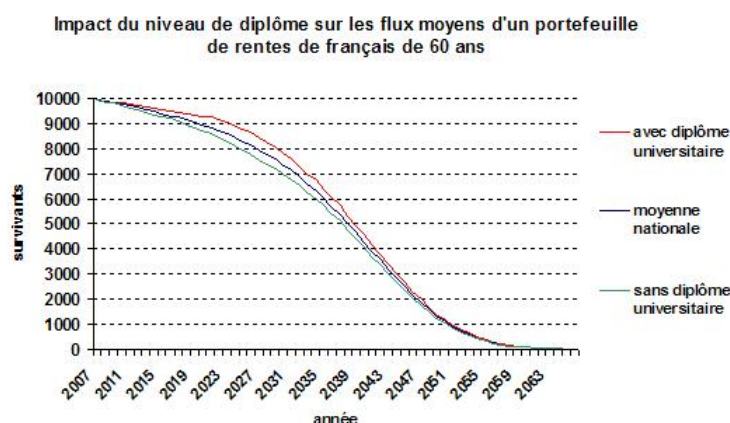


FIGURE 8.6 – Flux moyens d'un portefeuille constitué de 10000 hommes de 60 ans qui ont tous un niveau de diplôme donné

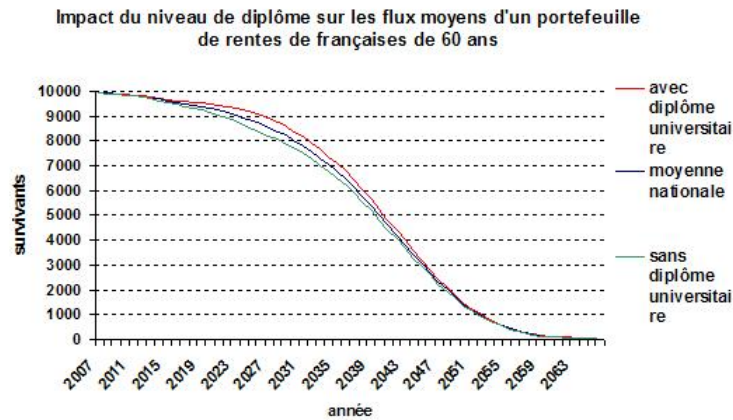


FIGURE 8.7 – Flux moyens d'un portefeuille constitué de 10000 femmes de 60 ans qui ont tous un niveau de diplôme donné

On remarque que le niveau de diplôme a une influence significative sur la mortalité aux âges élevés. De plus, cette influence est d'un ordre comparable pour l'homme et la femme. La variable de l'éducation est donc un moyen intéressant de prendre en compte le statut socioéconomique.

Influence du statut matrimonial

En assurance, il existe des différences significatives de rentes selon le statut matrimonial de l'individu. On sait que cette variable a une influence importante aux âges élevés dans la mesure où le mariage est protecteur en comparaison du divorce et du veuvage.

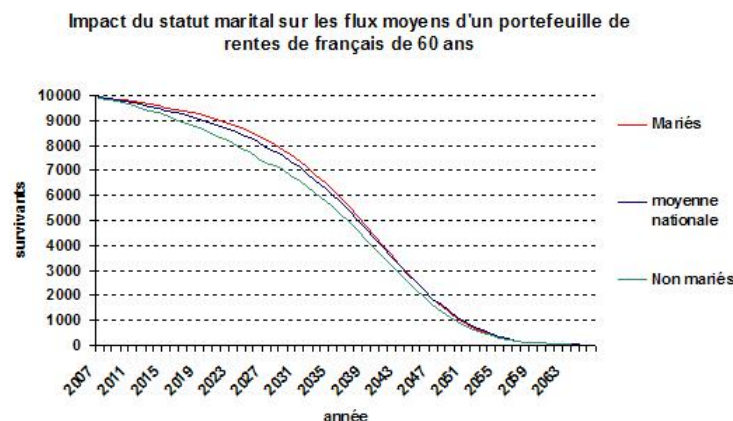


FIGURE 8.8 – Flux moyens d'un portefeuille constitué de 10000 hommes de 60 ans qui ont tous un statut matrimonial donné

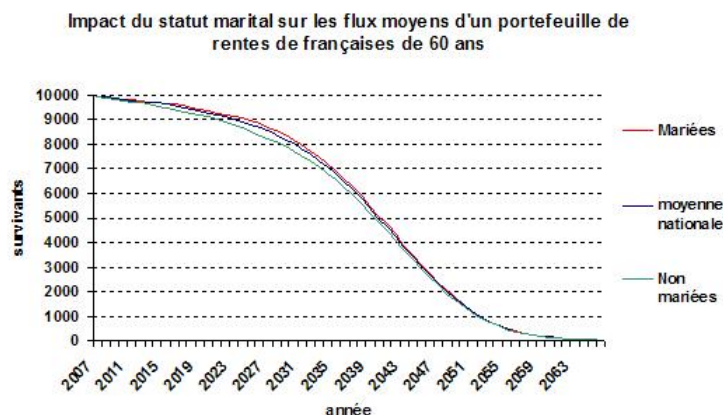


FIGURE 8.9 – Flux moyens d'un portefeuille constitué de 10000 femmes de 60 ans qui ont tous un statut matrimonial donné

Ces graphiques révèlent une influence majeure du statut matrimonial sur la valorisation d'un portefeuille de rentes, en particulier chez les hommes où les veufs et les divorcés sont fortement désavantagés par rapport aux hommes mariés. Elle apparaît être la caractéristique la plus discriminante en terme de longévité. Cependant, on a considéré que le statut matrimonial était une variable fixe. Il faut tenir compte de l'évolution du statut matrimonial dans le temps, ce qui modifiera les séries de flux à venir.

En tenant compte de l'évolution du statut matrimonial au cours de la vie de l'assuré, les écarts de mortalité seraient réduits. Il existe des taux de divorce en fonction de la durée de mariage par exemple (voir Chapitre 7, Section 2.4), et il est possible de modéliser l'occurrence des divorces.

Le veuvage peut être pris en compte de plusieurs manières. Une première idée est d'évaluer une probabilité individuelle de veuvage mais cette évaluation est difficilement réalisable du fait de la dépendance importante de cette probabilité par rapport aux caractéristiques du conjoint. Aussi, une autre manière naturelle de modéliser le veuvage est de décrire les caractéristiques du conjoint de l'assuré et de simuler son évolution simultanée dans le temps. Lorsque le conjoint décède, l'assuré devient veuf. Cette modélisation renseigne sur l'évolution du statut matrimonial de l'assuré.

Bilan sur le risque de base

Ces modèles d'évolution par âge et par traits permettent de valoriser les portefeuilles de rentes et d'estimer le risque de base résultant de l'hétérogénéité du portefeuille. On s'est limité à un portefeuille de personnes âgées de 60 ans et de caractéristiques données. On

aurait pu avoir un portefeuille totalement hétérogène avec différents âges et différentes caractéristiques individuelles, mais cela n'aurait pas été pédagogique à moins d'avoir accès aux données d'un portefeuille de rentes réel.

De plus, grâce à cette modélisation microscopique, il est aussi possible de distinguer chaque individu et de considérer que certains individus ont des poids plus importants que d'autres sur la valeur du portefeuille. En effet, en décrivant le montant de la cotisation de chaque individu, l'impact du décès d'un individu sur les flux rentes du portefeuille va dépendre de son nominal de rente.

8.3 Introduction à la couverture de produits d'assurance vie

Nous allons essayer de comprendre les différents risques mis en jeu et les objectifs des assureurs concernant la gestion de ces risques. Nous verrons qu'il est difficile de mettre en place des produits de couverture du risque pur de longévité. Nous analyserons aussi les méthodes de transfert de risque de taux des produits d'assurance vie.

8.3.1 Compréhension et gestion des risques en assurance vie

Les produits d'assurance vie ont une composante longévité et une composante financière dont le risque de taux d'intérêt. Afin de bien comprendre la spécificité des risques d'un produit d'assurance vie, on considère l'exemple d'un portefeuille de rentes de femmes françaises qui ont 60 ans pour lequel l'assurance doit payer à chaque assurée une rente durant sa vie.

La compagnie d'assurance est tout d'abord exposée au risque de longévité dans la mesure où les flux dépendent du nombre de survivants. Une des caractéristiques principales de ce risque est son horizon très long. En effet, les assurés de 60 ans ont une espérance de vie de l'ordre de 25 ans et la durée de vie de ce produit financier est de l'ordre de 45 ans, temps au delà duquel la quasi totalité des assurés sont décédés. Le montant des flux à verser diminue de moitié au bout d'une vingtaine d'années et baisse significativement à partir de 30 ou 40 ans. Néanmoins, les flux du produit restent tout de même importants et le risque long terme de longévité est par conséquent difficile à couvrir.

Gestion du risque de longévité

Les compagnies d'assurance doivent gérer ce risque de longévité et des produits ont été mis en place pour le transfert de risque. Il existe un grand nombre de produits permettant la gestion du risque de mortalité et du risque de longévité comme les swaps de longévité, les q-Forward, les longevity bond [6, 26, 2]. Cependant, les produits permettant de couvrir le risque de longévité n'ont pas connu un franc succès du fait de l'absence d'indice de référence

de longévité.

Il existe plusieurs causes majeures expliquant les difficultés à la mise en place d'un marché sur la longévité. Nous allons en recenser quelques unes :

- La réticence des investisseurs à couvrir un risque long terme. En effet, ces contrats long terme ne sont pas très attractifs dans la mesure où les bilans annuels ne sont pas améliorés avec ce type de transaction. Cependant, avec la nouvelle problématique du contrôle des risques, ces contrats pourraient devenir nécessaires.
- L'absence d'indice de longévité ne permet pas aux différents intervenants d'avoir une référence commune. Les contrats ne peuvent être mis en place sans estimation d'amélioration moyenne de longévité qui satisfait l'ensemble des acteurs de marché. Un indice national produit par le gouvernement pourrait être une solution intéressante [1].
- La modélisation de l'évolution de la longévité n'est pas universelle et les banques n'ont ni l'information ni la modélisation suffisante pour supporter un risque de longévité.
- Il n'existe pas de méthode standard de valorisation des produits sur longévité ce qui engendre des désaccords sur le prix du produit. L'évaluation ne peut se faire sous la probabilité risque neutre dans la mesure où les produits ne sont pas liquides et que le marché est incomplet. L'évaluation pourrait se faire par prix d'indifférence ou par d'autres méthodes classiques en marché incomplet (voir Chapitre 4).
- Les contrats proposés ne permettent pas à l'assureur de gérer intégralement son risque réel de longévité dans la mesure où le produit n'est pas adapté à son risque spécifique. La plupart du temps, le contrat permet de couvrir grossièrement le risque systémique alors que le risque spécifique n'est pas pris en compte.
- Le risque de défaut du vendeur de protection doit être pris en compte étant donné que l'horizon de ce genre de produit est lointain. La période de crise actuelle ne favorise pas l'émergence des produits de couverture de longévité.

On a vu que le risque de longévité pur est difficile à gérer dans les circonstances actuelles et que les marchés sur longévité ne sont pas encore mis en place. On s'intéresse dès lors au risque de taux d'intérêt qui est couplé au risque de longévité.

Risque de taux d'intérêt

Un produit d'assurance vie nécessite la gestion d'un risque de longévité important. Les compagnies d'assurance s'exposent aussi à un risque financier dans la mesure où elles investissent les cotisations perçues dans les marchés financiers, principalement dans le marché des taux d'intérêt qui rassemble plus de 80% de leurs investissements. L'évolution des taux d'intérêt est donc un risque très important que l'assurance se doit de couvrir.

L'assureur effectue certaines opérations qui lui permettent de gérer son risque de taux d'intérêt sans pour autant le couvrir dynamiquement. En effet, il essaie généralement de gérer de manière statique le risque de taux de son portefeuille de rentes en achetant un panier de swaps de taux de maturités différentes. La quantité de swap aux différentes maturités est choisie de manière à reproduire la courbe d'évolution du portefeuille estimée par l'assureur.

Avec cette stratégie statique, l'assureur pourrait échanger un taux fixe contre un taux variable sur l'échéancier des rentes. En effet, la durée de vie des assurés du portefeuille a une distribution centrée sur l'espérance de vie. L'assureur pourrait utiliser cette stratégie afin de réduire son risque de taux statique lié aux rentes. Cependant, cette technique pratique est assez grossière dans la mesure où l'ensemble des maturités de swap ne permettent pas de suivre précisément les flux du portefeuille. De plus, ce type d'opérations ne permet pas de gérer les risques dynamiquement puisqu'il ne considère que l'information qu'il a aujourd'hui. Dans le futur, les conditions auront changé et la stratégie que l'assureur a adoptée aujourd'hui ne sera plus adaptée. Il faut donc trouver un produit qui permettrait d'avoir un point de vue dynamique concernant le risque de longévité et le risque de taux.

La compagnie d'assurance peut aussi chercher à couvrir ce risque "forward" de taux d'intérêt sur les flux restants à verser. En effet, si on se place dans 10 ans par exemple, l'assureur devra gérer le risque de taux pour les flux de rentes dans le futur des personnes survivantes qui auront 70 ans dans 10 ans. L'assurance est souvent prête à gérer un risque futur de longévité mais aimerait se débarrasser du risque futur de taux d'intérêt. Il faudrait donc définir un produit qui ne soit pas trop affecté par le risque de longévité des individus et qui permettrait à la compagnie d'assurance de transférer son risque de taux à une date ultérieure. On verra que cela sera possible grâce à un produit qui permet à l'assureur de couvrir dynamiquement et quasi parfaitement son risque de taux et permettant à la banque de vendre un produit qu'elle sait à priori gérer.

8.3.2 Idées de produits de transfert du risque de taux

Nous présentons deux premières idées de produits qui pourraient permettre une gestion dynamique du risque de taux des produits d'assurance vie. Ces produits ne sont pas bien adaptés pour un transfert intéressant du risque de taux d'intérêt mais leur description permet de comprendre les qualités nécessaires d'un produit efficace de transfert de risque de taux.

Swaptions sur longévité

Nous introduisons un premier produit hybride taux/longévité. Ce produit va permettre de comprendre différentes problématiques de modélisation et de rappeler les difficultés liées à l'évaluation de ce genre de produit.

Description du produit Si un assureur veut couvrir grossièrement le risque de taux moyen inhérent aux rentes des personnes de 70 ans dans 10 ans, il cherche à rentrer dans un produit de taux dont la maturité est liée à la longévité des individus. Ainsi, l'assureur pourrait vouloir acheter une swaption $Sw(10Y, 10Y, 10Y + f(e_{10}^{70}), 1Y)$ dont la maturité du swap sera une fonction de l'espérance de vie e_{10}^{70} des personnes de 70 ans dans 10 ans. Cette espérance de vie sera fixée dans 10 ans par des tables de mortalité ou un indice national : il n'existe pas d'indice national prospectif mais dans le futur, il sera possible de déduire l'espérance de vie e_{10}^{70} à partir des tables de mortalité dans 10 ans.

Ce contrat donne la possibilité dans le futur de rentrer dans un swap dont la maturité sera bien adaptée au risque qu'il cherche à couvrir. En effet, $e_{10}^{70} \sim 15Y$ sera bien évidemment différent de e_0^{60} et ce produit tiendra compte de l'évolution stochastique de la longévité. La fonction f permet d'ajuster la maturité au goût de l'assureur : il chercherait par exemple à avoir une maturité assez grande pour que l'ensemble des assurés soient presque tous décédés et il souhaiterait acheter une swaption sur une maturité supérieure à e_{10}^{70} . Il pourrait considérer par exemple $f(e_{10}^{70}) = e_{10}^{70} + k_{70}$ avec une valeur de k_{70} qui correspondra à l'estimation de l'âge à partir duquel une certaine proportion α des individus ($\alpha = 90\%$ par exemple) seront censés être décédés. Cela lui assurera de couvrir une grande partie du risque de taux de son portefeuille de rentes. Cette constante k_a (dépendant de l'âge des assurés) pourra par exemple être calculée par l'assureur sur la distribution actuelle des durées de vie résiduelles en supposant que les queues de la distribution ne vont pas beaucoup s'épaissir dans le temps. Cela sera un degré de liberté de l'assureur.

Modélisation du produit Ce produit a une composante taux et une composante longévité ce qui le rend assez complexe dans la mesure où l'acheteur et le vendeur de ces contrats doivent savoir modéliser les deux types de risques. La modélisation peut d'ailleurs considérer que les risques de taux et de longévité sont éventuellement corrélés et se révèlent très complexes pour ce type de produit.

La gestion du risque de longévité nécessite des données et un savoir-faire spécifique dont la banque ne dispose généralement pas. Par conséquent, il sera difficile de mettre d'accord les deux intervenants sur ce type de produit sans indice de référence de longévité.

Évaluation du produit Concernant l'évaluation du produit, elle est aussi compliquée dans la mesure où elle va dépendre d'un modèle de longévité et qu'elle se fera en marché incomplet. En effet, le produit mélange la modélisation des taux d'intérêt à celle de la longévité ce qui rend son évaluation assez complexe. Il n'est pas sûr que les banques soient assez renseignées sur la longévité pour traiter sereinement ce type de produit. De plus, les assurances ne valideront pas l'évaluation effectuée par la banque. Pour que l'assurance et les banques puissent s'engager dans un contrat, il faut qu'elles disposent de la même infor-

mation concernant les modèles de taux et les modèles de longévité, ce qui apparaît assez utopique.

Ceci étant dit, on peut tout de même chercher à évaluer ce produit qui est tout de même adapté à la couverture partielle du risque de taux lié à certains contrats sur longévité. Il est à noter que l'évaluation en probabilité risque neutre (ou forward neutre) ne sera pas bien adapté lorsque la maturité T_N dépendra de la longévité. En effet, le marché étant incomplet et très peu liquide, l'évaluation risque neutre ne peut convenir à ce cadre. Il faut par conséquent utiliser des méthodes d'évaluation en marché incomplet pour ce produit.

On considère les filtrations à la date 0 liées à l'information disponible sur la courbe de taux \mathcal{F}_0 et à l'information concernant les tables de mortalité disponibles \mathcal{G}_0 . On pose $\mathcal{H}_0 = \mathcal{F}_0 \vee \mathcal{G}_0$. Sachant cette information ainsi que les flux de la jambe correspondant à $F = (K - S_T(T_0, f(e_T^a), \delta))^+ \sum_{i=1}^N B(T, T_i)$, on cherche à l'évaluer dans un marché incomplet tenant compte de l'information sur la longévité et sur les taux d'intérêt. Il existe plusieurs types de méthodes d'évaluation dans ce type de marché comme les méthodes d'évaluation statique, l'évaluation par prix d'indifférence [21], le calcul de prix d'équilibre pour de petites transactions [12]. Il faudra par conséquent essayer de travailler sur la longévité et les taux d'intérêt en probabilité historique.

Maintenant que l'on a relevé les différents enjeux liés à la modélisation et à l'évaluation de ce type de produit, il faut aussi qu'on explique pourquoi ce premier produit n'est pas adapté au transfert de risque entre une compagnie d'assurance et une banque.

Analyse pratique du produit Ce produit n'est pas adapté à la gestion du risque du taux d'intérêt car il ne tient pas compte de l'évolution des assurés encore vivants du portefeuille à qui il faut payer effectivement une rente. Chaque année, des individus décèdent et les nominaux des rentes diminuent dans le temps. Cette swaption permet un transfert partiel et grossier du risque de taux qui n'est pas spécifique au portefeuille.

Ce produit n'est pas très intéressant pour l'assurance n'étant pas directement adapté à sa problématique. De plus, la banque n'est elle non plus pas tout à fait satisfaite. En effet, l'information liée à la longévité est contenue dans la maturité du swap et le produit contient un risque de longévité important. Si la banque devait vendre ce produit, il lui faudrait une modélisation fine sur la longévité afin d'estimer au mieux e_{10}^{70} en date 0. Bien que ce produit ne soit pas vraiment adapté aux besoins réels de l'assureur, il présente l'intérêt de remarquer que la modélisation et l'évaluation de produits d'assurance vie peuvent facilement devenir compliquées.

Étude de la Longevity Maturity Chooser Swaption

On décrit ici rapidement un deuxième produit qu'on appelle la *Longevity Maturity Chooser Swaption* (LMCS) répondant à la plupart des inconvénients de la swaption sur longévité. Ces produits donnent des pistes pour trouver un produit adapté à la problématique du transfert de risque de taux.

Description du produit La banque pourrait proposer un produit assez intéressant qui donnerait le droit à l'acheteur du produit de rentrer dans une swaption dont la maturité du swap est comprise dans un intervalle $[T_{min}, T_{max}]$ et dont les deux bornes seront fixées par l'assureur au début du contrat. Cet intervalle de maturité va permettre à l'assureur de choisir la maturité du swap qui dépendra a priori de la longévité des individus. En effet, avec un modèle de longévité, l'assureur va estimer la distribution de l'espérance de vie dans le futur et/ou la durée de vie résiduelle des assurés dans le futur, et choisira la maturité de swap qu'il considérera comme la plus adaptée.

Plaçons nous dans le cas où l'assureur souhaite rentrer dans un swap de maturité la plus proche de e_{10}^{70} car il souhaite gérer son risque de taux dans 10 ans pour cette période. Grâce à ses modèles de longévité, il peut estimer aujourd'hui e_{10}^{70} si son portefeuille est constitué de rentes pour des assurés qui ont actuellement 60 ans et qui en auront 70 dans 10 ans. Il peut aussi estimer des maturités extrêmes qui correspondront généralement aux bornes d'un intervalle de confiance sur l'estimation de cette espérance de vie. Autrement dit, ces maturités représentent généralement deux quantiles de la distribution de la durée de vie dans le futur. L'assurance souhaiterait avoir le droit de rentrer dans 10 ans dans un swap dont la maturité est une maturité quelconque comprise entre ces deux maturités. En pratique, la maturité idéale sera l'espérance de vie e_{10}^{70} mais il faut que l'intervalle $[T_{min}, T_{max}]$ soit assez large pour avoir une forte chance de la contenir. Après avoir déterminé ces bornes de confiance, l'assureur achète à la banque ce produit qui donnant le droit de rentrer dans un swap dont la maturité pourra être choisie entre ces deux bornes. L'intervalle des maturités comprises entre T_{min} et T_{max} permettent à l'assureur de tenir compte de l'évolution de la longévité afin de prédire les flux de rentes à venir.

L'intérêt de ce produit, c'est que l'assureur gère totalement le risque longévité et que la banque lui offre la possibilité de gérer le risque de taux inhérent au risque de longévité. L'évaluation du produit ne dépend d'aucun modèle de longévité mais seulement d'un modèle de taux d'intérêt. Cela facilitera le dialogue entre les deux acteurs de marchés et simplifiera grandement les négociations. Le prix du produit est fixé par la banque (pour toutes les valeurs de T_{min} et de T_{max}) et sera facilement accessible dans la mesure où la banque n'aura pas besoin d'être spécialisée en longévité pour traiter ce produit. L'assureur estime les deux bornes qui lui convient pour son portefeuille et propose d'acheter ce produit à la

banque dont le prix est indépendant d'un modèle de longévité.

Évaluation du produit De plus, l'évaluation et l'étude de ce produit sont assez intéressantes car le prix du produit ne dépend pas trivialement des bornes T_{min} et T_{max} . Ces bornes pourront être ajustées de manière à ce que le prix soit proche d'un prix attendu. L'évaluation peut se faire de manière risque neutre (les maturités sont liquides) et fournit l'évaluation suivante :

$$\Pi(T, T_0, T_{min}, T_{max}, \delta) = B(0, T) \mathbb{E}^{Q_T} \left[\max_{T_{min} \leq T_N \leq T_{max}} \left\{ (K - S_T(T_0, T_N, \delta))^+ \sum_{i=1}^N B(T, T_i) \right\} \right].$$

Étude qualitative En effet, il existe deux effets à étudier sur l'évaluation du produit. Le premier effet est qu'à prix d'exercice fixé, il serait intéressant pour l'acheteur du produit de prendre le taux de swap le plus faible pour que le flux $(K - S_T(T_0, T_N, \delta))^+$ soit le plus grand possible. En effet, lorsque la courbe des taux initiale est croissante, la courbe des taux de swap est "qualitativement" croissante et par conséquent il faudrait choisir le swap de maturité la plus courte. Cette considération sur l'équivalence des comportements de courbes pour différents taux (actualisations, TEC, taux de swap,...) est réellement vérifiée en pratique.

Le deuxième effet est que l'acheteur du produit cherche à recevoir un flux le plus longtemps possible et par conséquent préférera la maturité du swap la plus longue. Ces deux effets vont plus ou moins se compenser et vont donner une étude très intéressante pour ce produit.

Décrivons un exemple intuitif expliquant l'aspect non trivial du produit. Considérons que la courbe de taux de swap dans 10 ans possède 2 taux de swap $Sw(10Y, 10Y, 20Y) = 3.5\%$ et $Sw(10Y, 10Y, 30Y) = 4\%$ et que le taux fixe est à 4.5%. Ces choix de maturités de swap correspondent au cas où l'assureur estime que l'espérance de vie à 70 ans dans 10 ans sera comprise entre $T_{min} = 10Y$ et $T_{max} = 20Y$ (avec une certaine marge) et qu'il veut avoir la possibilité dans 10 ans de rentrer dans un swap avec une maturité comprise entre ces valeurs. Simplifions et disons que l'assureur a le choix entre rentrer dans le swap de maturité 10 ans et le swap de maturité 20Y (en fait avec le produit décrit précédemment, il pourrait rentrer dans n'importe quelle maturité intermédiaire). Qualitativement, le swap numéro 1 lui fournira un gain de 1% pendant 10Y soit environ 10% (si la courbe des taux est plate). Le swap numéro 2 va fournir un gain de 0.5% pendant 20 ans soit environ 10%. Ces deux stratégies sont discutables et il n'est pas évident de savoir dans quel swap l'acheteur aura intérêt à rentrer. Par conséquent, des tests sont nécessaires pour déterminer quel swap est le plus cher vu d'aujourd'hui et comment évaluer le prix de ce type de produit. Cela sert aussi à justifier le fait que le prix ne va pas vraiment diverger en faisant bouger les bornes des paramètres car dans ce cas (avec des taux assez proches de la réalité), il n'est pas évident de savoir quel swap coûtera le plus cher.

Remarque 8.2. *Une étude quantitative de ce produit a été faite mais nous ne la présentons pas dans ces travaux puisqu'on préfère mettre l'accent sur le produit principal de la section suivante.*

Analyse du produit Ce produit peut être intéressant étant donné que c'est un produit pur de taux et que l'évaluation peut se faire en probabilité risque neutre avec une liquidité suffisante. De plus, il a l'avantage de pouvoir devenir un produit liquide dans la mesure où les banques pourraient donner un prix de marché (systématique) à ce produit par chaque plage de maturités de swap $[T_{min}, T_{max}]$ et pour chaque maturité de swaption.

Cependant, ce produit ne règle pas le problème majeur du fait qu'il ne permet à la compagnie d'assurance de faire qu'un transfert grossier de son risque de taux. En effet, l'assurance aimerait pouvoir transférer le risque de taux spécifique de son portefeuille de rentes et ce produit de taux ne tient pas compte de l'évolution des assurés dans le temps.

Qualités requises du produit de transfert de risque

Nous avons vu qu'un produit qui dépend de la dynamique de longévité pose de nombreux problèmes et ne pourra pas être mis en place facilement. Ce qui nous amène à considérer un produit de taux. Il faut cependant qu'il soit bien adapté aux besoins spécifiques de l'assurance ce qui n'était pas le cas de la "Longevity Maturity Chooser Swaption".

Par conséquent, nous allons chercher à mettre en place un produit purement financier qui est lié indirectement à la longévité. Ce produit va permettre de décorréliser les risques de longévité et de taux d'intérêt d'un autre produit. La compagnie d'assurance qui possède des portefeuilles de rentes pourra estimer son propre risque de longévité et pourra transférer l'intégralité du risque de taux d'intérêt sous-jacent à une institution financière tout en conservant le risque de longévité.

Tout d'abord, nous allons essayer de mettre en place un produit purement financier qui ne pose pas de problèmes majeurs concernant la modélisation ou l'évaluation. De plus, l'absence d'indice de référence suggère la structuration d'un produit adapté aux besoins de l'assurance. Dans un souci d'efficacité du produit, la méthode de valorisation doit être générale et doit pouvoir s'appliquer à n'importe quelle situation.

Une idée simple de produit est un produit pur de taux qui ne dépend pas directement de la dynamique de la longévité mais qui en dépend indirectement. Ce type de produit limite les désaccords sur la modélisation de longévité et permet à la banque de ne pas avoir à gérer le risque de longévité.

8.4 Longevity Nominal Chooser Swaption (LNCS)

Pour les produits décrits dans la partie précédente, nous n'avions pas considéré simultanément le point de vue de l'assurance et de la banque. La "Longevity Maturity Chooser Swaption" conviendra certainement à la banque mais ne répondra pas aux besoins spécifiques de l'assurance. En effet, la compagnie assurance pourra couvrir "partiellement" son risque de taux en achetant une swaption dont la maturité de swap sera adaptée à ses besoins mais le fait que le nominal des flux du swap soit constant ne traduit pas le cas concret d'un portefeuille de rentes pour lequel le nominal va évoluer avec le temps.

8.4.1 Introduction au produit

Nous allons à nouveau introduire un produit de taux afin de faciliter la modélisation et les méthodes d'évaluation. L'information sur la longévité des individus sera captée par une série de nominaux variables. Nous allons donc considérer le cas d'une swaption plus compliquée dans lequel l'acheteur a la possibilité de rentrer dans un swap dont les nominaux varient chaque année. En effet, l'assureur utilise son modèle de longévité et peut estimer des évolutions extrêmes de survie des assurés de son portefeuille de rentes. L'évolution des assurés dans le futur peut être associée à des séries de nominaux.

Description du risque de longévité

On considère un portefeuille réel d'hommes français en 2009. L'information disponible est la distribution en âge des assurés qui est décrite sur le graphique suivant (voir Annexe A pour les données) :

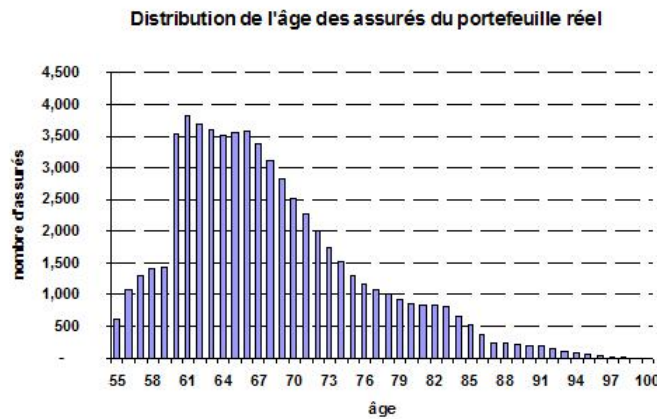


FIGURE 8.10 – Répartition des âges des assurés du portefeuille

Connaissant cette population initiale constituée de 62482 en 2009, il est possible de simuler

des scénarios d'évolution de ces assurés dans le temps. On analyse le scénario médian et les scénarios extrêmes en renseignant les quantiles à 5% et 95%. L'assureur estime que la courbe des flux de son portefeuille de rentes sera a priori comprise entre ces deux bornes.

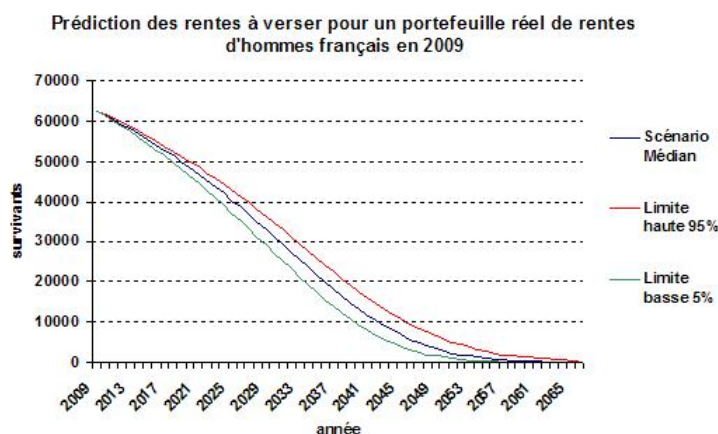


FIGURE 8.11 – Exemple de scénarios extrêmes de survie des assurés du portefeuille

Ces courbes sont renormalisées dans le sens où nous avons supposé une rente de 1 euro par assuré. Si chaque assuré reçoit en moyenne une rente de 1000 euros par an, cela correspond à un montant d'environ 60 millions d'euros la première année et ces montants diminueront au fur et à mesure que les assurés du portefeuille décéderont. Les montants pouvant être en jeu peuvent être colossaux et il est donc important pour l'assureur de couvrir le risque de taux sur les flux à venir en tenant compte de l'évolution des assurés.

Analyse des projections

Historiquement, les individus gagnent chaque année environ 3 mois d'espérance de vie ce qui revient aussi à considérer qu'en 10 ans, les individus gagnent environ 2 ans et demi d'espérance de vie. On se rend compte de l'intérêt d'utiliser des tables prospectives (provenant de modèle de mortalité stochastiques) plutôt que des tables de mortalité statiques dans la mesure où les tables prospectives tiennent compte de l'amélioration de la longévité dans le temps. On dispose à présent des deux courbes d'évolution avec le modèle de dynamique de population. Les deux bornes sont assez rapprochées les premières années mais l'écart se creuse rapidement. En date initiale, l'assureur est soumis à un risque de taux sur une série de flux incertaine. On met en place un produit qui va permettre à l'assurance de couvrir son risque de taux sur une série de flux comprise entre la courbe inférieure $\{N_t^-, 0 \leq t \leq T_M\}$ et la courbe supérieure $\{N_t^+, 0 \leq t \leq T_M\}$. On note T_M la maturité maximale sur laquelle on décrit les séries de nominaux.

La compagnie d'assurance cherche à couvrir dans X années le risque de taux lié à son portefeuille de rentes. Dans cet exemple, on va considérer que $X = 10$ ans et l'assureur peut tenir compte de l'évolution de la longévité et des taux d'intérêt pendant les dix premières années. Plus X est grand, meilleure sera l'estimation des flux de rentes restants puisque l'assureur va capter l'évolution de la longévité nationale ainsi que l'évolution de son portefeuille sur une période plus longue.

Dans $X = 10$ ans, l'assureur recense le nombre de personnes encore vivantes N_T de son portefeuille et va chercher un nombre α_T fixé en T tel que $N_T = \alpha_T N_T^- + (1 - \alpha_T) N_T^+$. Cet α_T va lui donner une meilleure estimation de la série de flux restant que l'estimation qu'il aurait pu faire en date initiale puisqu'il tient compte de l'évolution du portefeuille pendant les X premières années. En date 0, cette méthode est une bonne manière d'estimer les flux futurs sachant l'information qui sera disponible en date T . La série des flux à venir pourra être considérée proche de la série de nominaux donnés par $N_t = \alpha_T N_t^- + (1 - \alpha_T) N_t^+$ pour $0 \leq t \leq T_M$ (voir Figure 8.11).

Exemples de choix du paramètre α_T

Prenons l'exemple suivant du portefeuille d'assurés en 2009. Les courbes extrêmes, provenant du modèle de longévité, révèlent que le nombre de survivants dans 10 ans sera normalement compris entre 48048 et 51223 individus avec un scénario médian de 49888 personnes. Il existe plusieurs méthodes pour estimer le paramètre α_T .

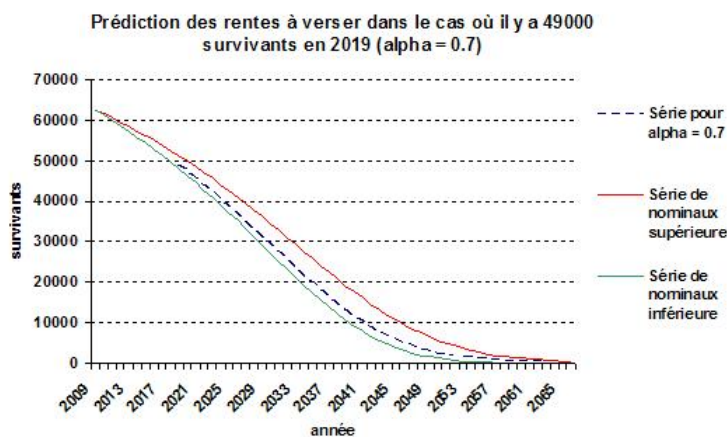


FIGURE 8.12 – Choix du paramètre α_T en 2019

Une méthode naturelle consisterait à regarder le nombre de survivants et d'adapter le α_T en fonction de cela. Si en pratique il y a 49000 personnes vivantes en $T = 2019$, l'assureur calculera en date T un $\alpha_T = \frac{N_T^+ - N_T}{N_T^+ - N_T^-} = \frac{51223 - 49000}{51223 - 48049} \sim 0.7$. L'évolution des survivants dans

les années à venir peut être estimée par : $N_t = 0.7N_t^- + 0.3N_t^+$. L'assureur va donc chercher à se couvrir contre un mouvement de taux qui auront une influence sur les rentes à verser (voir Figure 8.12).

On étudie le nombre de survivants en $T = 2029$ qui est censé être compris entre 29126 et 36179 avec un scénario médian à 33009 personnes. Considérons que la mortalité a été plus élevée ou que la longévité ne s'est pas autant améliorée que ce que prédisait la tendance et qu'il reste plutôt 34750 survivantes en 2029. Il est possible de calculer un α_T qui fournira un nouveau scénario tenant compte des événements passés jusqu'en 2029. On calcule le $\alpha_T = \frac{N_T^+ - N_T}{N_T^+ - N_T^-} = \frac{36179 - 34750}{36179 - 29126} \sim 0.2$ et on estime la suite de flux à venir comme sur la Figure 8.13.

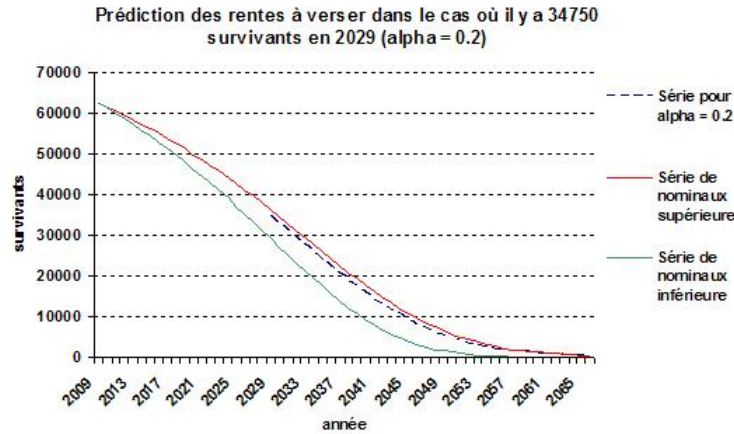


FIGURE 8.13 – Choix du paramètre α_T en 2029

Il est important de noter qu'il existe d'autres choix pour le α_T . Par exemple, entre la date initiale et la date d'exercice du contrat (ici 10 ans), il pourrait y avoir une rupture de tendance concernant l'évolution de la mortalité des individus. Cette rupture de tendance pourrait être due à une découverte technologique ou médicale. Les modèles actuels (comme le modèle de Cairns ou Lee Carter) projettent une tendance mais ne tiennent pas compte des ruptures de tendance éventuelles pouvant apparaître à certains moments. Si l'assureur se rend compte qu'une rupture de pente a eu lieu, il pourra adapter la nouvelle série de flux à venir en choisissant un α_T "mieux adapté". Cela lui permet de gérer son risque de longévité bien que le modèle ne pouvait prédire cette évolution en date initiale. Le paramètre α_T pourrait aider à tenir compte de ces phénomènes qu'on ne prend pas en compte dans les projections de mortalité.

Remarque 8.3. Cette remarque sur la rupture de tendance laisse entrevoir un autre produit

qui s'intéresserait les tendances d'amélioration de longévité. Par exemple, on pourrait penser à un produit qui considère que la tendance sera comprise entre deux valeurs dont on estime les valeurs extrêmes. Cela laisserait le choix à l'assureur dans X années de choisir une nouvelle tendance d'évolution plus adaptée dans le futur.

Hétérogénéité du portefeuille

Le modèle de dynamique de population peut être utilisé afin de générer des scénarios d'évolution d'un portefeuille d'assurés. Dans l'exemple traité, on ne renseignait que la distribution de l'âge des assurés qui étaient tous de sexe masculin cependant on peut facilement considérer un portefeuille hétérogène concret. En effet, le modèle étant un modèle microscopique à l'échelle de l'individu, chaque assuré est caractérisé par son âge, son sexe, sa catégorie sociale et/ou d'autres traits.

Il est aussi possible d'associer à chaque individu un nominal individuel de rente et non un nominal moyen comme c'est souvent le cas. Cela est très important dans la mesure où certains individus ont des rentes bien plus importantes que d'autres, et que leur décès ou leur survie aura un impact important sur la valorisation du portefeuille. Il est très difficile de traiter le cas d'un portefeuille inhomogène avec un modèle de mortalité classique pur car la modélisation n'est pas microscopique mais macroscopique. De ce point de vue le modèle de dynamique de population présente de nombreux avantages dont celui de générer des scénarios d'évolution pertinents (voir Chapitre 7) et qui tient compte de l'ensemble de l'information disponible.

Risque de taux d'intérêt

On s'intéresse à l'étude du risque de taux d'intérêt inhérent au portefeuille de rente. Des hypothèses simplificatrices sur les taux de rentes et les nominaux sont faites : on suppose pour simplifier que les portefeuilles sont classés par taux de rente c'est à dire que chaque assuré a le même taux de rente $k\%$. En réalité, il y a un taux de rente différent pour chaque assuré du portefeuille mais il est possible de calculer un taux de rente moyen du portefeuille comme cela est fait en pratique. Pour simplifier les notations, on suppose aussi que chaque assuré a le même nominal mais on verra plus tard qu'il est très facile de tenir compte d'un nominal individuel dans le modèle et le produit.

L'assurance est soumise au risque d'évolution des taux d'intérêt et il est assez complexe de comprendre son fonctionnement. La compagnie d'assurance a un rendement annuel qui dépend très fortement du niveau des taux d'intérêt dans la mesure où près de 80% de ses placements concernent le secteur des taux d'intérêt avec les dettes d'état, les obligations et autres produits de taux. L'assurance cherche à avoir un rendement supérieur au taux de rentes k qu'elle s'engage à payer aux assurés. Elle est donc fortement soumise à un risque

de taux d'intérêt.

Dans le scénario où les taux d'intérêt diminuent, son rendement global va lui aussi diminuer et il sera très difficile d'assurer le paiement des rentes au taux k . L'assurance est aussi soumise à un risque de montée des taux d'intérêt mais d'une manière tout à fait différente. En effet, en cas de montée des taux d'intérêt, les assurés ne voudront pas rester dans un contrat de rentes qui leur rapporterait un taux faible de $k\%$ alors qu'un placement à la banque pourrait leur procurer un meilleur rendement. Les assurances sont souvent forcées d'ajuster les taux de rente afin de conserver ses clients. Dans tous les cas, l'assurance cherche à couvrir son portefeuille d'un mouvement de taux d'intérêt que ce soit une augmentation ou une diminution des taux d'intérêt et nous allons essayer de mettre en place un produit qui permette de se couvrir contre ce genre de variations.

8.4.2 Rappels de certains concepts financiers

Avant de décrire le produit, il faut décrire certains concepts classiques dans les marchés financiers qui ne le sont pas forcément dans le monde de l'assurance. Dans un premier temps, nous allons décrire la notion de swap à nominal variable, la notion de swaption sur swap à nominal variable puis nous présentons en détails le produit de transfert de risque.

Swaps de taux d'intérêt

Le swap de taux d'intérêt de taux variables contre taux fixes est un contrat de gré à gré permettant l'échange d'intérêts d'un prêt ou dépôt notionnel à taux variable contre des intérêts à taux fixe. Les swaps sont des instruments très utilisés sur le marché des taux d'intérêt car ils permettent de se couvrir contre certains mouvements de marché.

Swaps standards Le marché des swaps standards (plain vanilla en anglais) contre taux IBOR constitue le deuxième plus important marché des taux d'intérêt à moyen et long terme, derrière celui des emprunts d'État et futures sur emprunts d'État. Le nominal est fixé en date initiale pour toute l'échéance du contrat. La série de flux associée au taux fixe est appelée jambe fixe et celle associée au taux variable, qui est généralement le taux Libor servant de référence principale sur le marché des swaps, est appelée jambe variable.

On considère un swap à la date T sur un échéancier débutant en T_0 et finissant en T_N d'un taux fixe k contre taux Libor. La valorisation standard des jambes fixes et des jambes variables est donnée par :

$$\begin{aligned} BF &= \sum_{i=1}^N k \delta B(T, T_i). \\ BV &= B(T, T_0) - B(T, T_N). \end{aligned}$$

On définit le taux de swap comme le taux fixe permettant l'égalisation des deux jambes ou qui annule la valeur du swap. Le taux de swap forward est fixé en T pour un swap commençant en T_0 de maturité T_N dont les paiements sont espacés de δ . Un simple raisonnement par absence d'opportunités d'arbitrage mène à :

$$S_T(T_0, T_N, \delta) = \frac{B(T, T_0) - B(T, T_N)}{\sum_{i=1}^N \delta B(T, T_i)}.$$

Swaps à nominal variable Un swap à nominal variable consiste en l'échange de deux jambes de flux dépendant d'une série de nominaux N_t . On considère un taux fixe k et le taux Libor $L(t, t + \delta)$ dont la valeur initiale est par exemple $L(0, \delta) = 5\%$ en date 0. La jambe fixe est constituée d'une série de flux de valeurs $\delta k N_t$ à la date t . La jambe variable est constituée d'une série de flux de valeurs $\delta L(t, t + \delta) N_t$ à la date t . Pour donner un exemple, on présente un exemple de trajectoire de l'évolution des taux Libor dans le temps Figure 8.14.

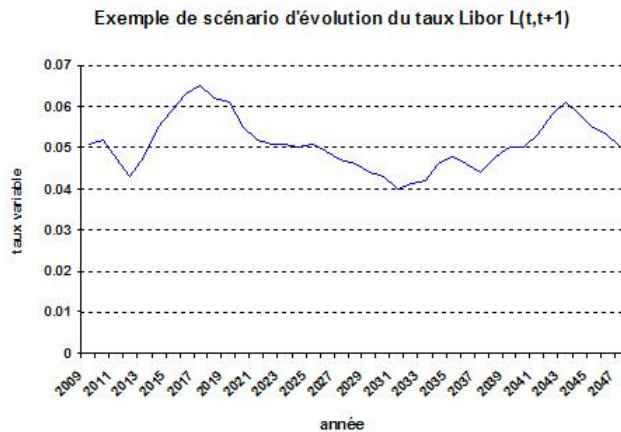


FIGURE 8.14 – Évolution du taux Libor $\delta L(t, t + \delta)$ en fonction de la date t

Cette trajectoire du taux Libor présente une baisse entre l'année 2009 et l'année 2012 puis une remontée conséquente des taux entre 2012 et 2020. Ce mouvement de taux va avoir un impact important sur la jambe variable du swap à nominal variable (voir Figure 8.16).

Un swap à nominal variable est un contrat gratuit à l'entrée qui permet d'échanger une jambe contre une autre. Pour cela, il faut trouver la valeur du taux fixe k de telle sorte que le prix de la jambe variable soit égal au prix de la jambe fixe. On se donne une série de nominaux N_t variables dans le temps et on suppose que $\delta = 1Y$ et que le taux fixe vaut $k = 5\%$ pour se fixer les idées. Avec ces différentes données, on peut obtenir un exemple de scénario pour les deux jambes d'un swap.

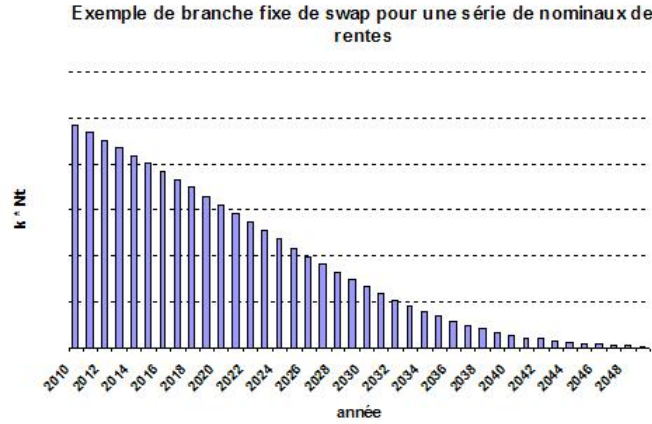


FIGURE 8.15 – Flux de la jambe fixe de swap à nominal variable à la date t : $\delta k N_t$

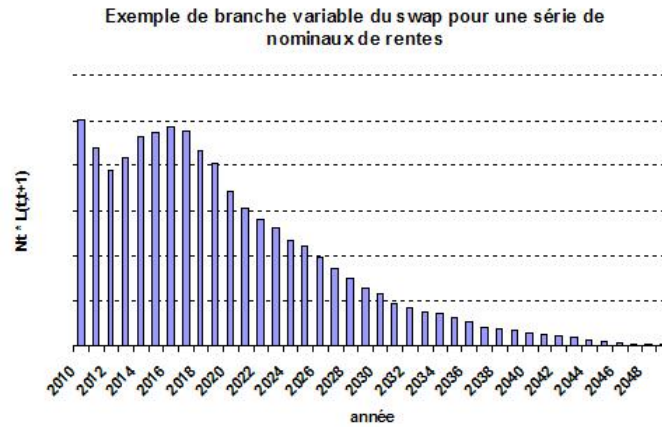


FIGURE 8.16 – Flux de la jambe variable de swap à nominal variable à la date t : $\delta L(t, t+\delta) N_t$

Pour ces deux graphiques, on a utilisé la courbe de nominaux provenant de la courbe médiane d'évolution de la survie des assurés. On remarque que l'évolution du taux court les premières années a un impact direct sur les flux de la jambe variable du swap (voir Figure 8.14 et Figure 8.16).

Le taux k permettant l'égalisation des deux jambes est appelé taux de swap forward à nominal variable $SV_T(T_0, T_N, \delta, N_t)$. La suite de nominaux $\{N_t\}$ étant fixée à la date t jusqu'à la date T_N et la jambe fixe s'évalue de la manière suivante :

$$BF = SV_T(T_0, T_N, \delta, N_t) \delta \sum_{i=1}^N N_{T_i} B(T, T_i).$$

Quant à la jambe variable, on la calcule facilement par l'approche du portefeuille d'obligations. En effet, une obligation de nominal N_1 à la date T_0 rapporterait la série de flux $\delta L(t, t + \delta)N_1$ en chaque date $t + \delta$ jusque la date T_N et fournit en date T_N le nominal N_1 . La valeur de cette obligation est 1 en date T_0 et vaut $B(T, T_0)$ en date T . Ainsi, cette jambe variable correspond aussi à une série d'obligations $(O_i)_{1 \leq i \leq N}$ commençant aux dates T_{i-1} et de nominaux $N_{T_i} - N_{T_{i-1}}$ et il ne faut pas oublier de retrancher les nominaux de chacune des obligations en date finale. Finalement, l'évaluation de cette jambe est donnée par :

$$\begin{aligned} BV &= N_{T_1} (B(T, T_0) - B(T, T_N)) \\ &+ (N_{T_2} - N_{T_1}) (B(T, T_1) - B(T, T_N)) \\ &+ \dots \\ &+ (N_{T_N} - N_{T_{N-1}}) (B(T, T_{N-1}) - B(T, T_N)). \end{aligned}$$

Cela peut se réécrire sous la forme :

$$BV = N_{T_1} (B(T, T_0) - B(T, T_N)) + \sum_{i=2}^N (N_{T_i} - N_{T_{i-1}}) (B(T, T_{i-1}) - B(T, T_N)).$$

En considérant la convention d'écriture "abusive" $N_{T_0} = 0$, on peut encore l'écrire sous la forme simplifiée suivante :

$$BV = \sum_{i=1}^N (N_{T_i} - N_{T_{i-1}}) (B(T, T_{i-1}) - B(T, T_N)).$$

Par conséquent, en égalisant la valeur des deux jambes du swap BV et BF pour une série de nominaux donnée, on peut calculer le taux forward de swap à nominal variable $SV_T(T_0, T_N, \delta, N)$ associé à une série de nominaux $\{N_t\}$:

$$SV_T(T_0, T_N, \delta, N_t) = \frac{\sum_{i=1}^N (N_{T_i} - N_{T_{i-1}}) (B(T, T_{i-1}) - B(T, T_N))}{\delta \sum_{i=1}^N N_{T_i} B(T, T_i)}.$$

Swaption sur swap de taux d'intérêt

Une swaption est une option sur swap de taux d'intérêt et représente pour celui qui la détient le droit de mettre en place à une échéance donnée, un swap dont les caractéristiques sont prédéterminées dans le contrat

Swaption standard On considère une swaption de maturité T sur un swap de date de fixing T_0 et une date de fin de swap T_N dont les flux sont séparés par un intervalle de temps δ . Le prix de cette swaption sera noté $Sw(T, T_0, T_N, \delta)$.

On distingue deux catégories de swaptions. Les swaptions acheteuses correspondent à des options d'achat qui permet de rentrer dans des swaps où l'acheteur reçoit le taux variable et paye le taux fixe K . Les swaptions receveuses correspondent à des options de ventes où l'acheteur peut rentrer dans un swap en recevant le taux fixe K et qui paye en contrepartie le taux variable. Par conséquent, le prix d'une swaption receveuse est donné par :

$$\begin{aligned} Sw(T, T_0, T_N, \delta) &= B(0, T) \mathbb{E}^{Q_T} \left[(K - S_T(T_0, T_N, \delta))^+ \sum_{i=1}^N \delta B(T, T_i) \right] \\ &= B(0, T) \mathbb{E}^{Q_T} \left[\left(K \sum_{i=1}^N \delta B(T, T_i) - B(T, T_0) + B(T, T_N) \right)^+ \right]. \end{aligned}$$

Swaption sur swap à nominal variable Après avoir défini un swap indexé sur la série de nominaux N_t , on peut définir une swaption qui donne le droit de rentrer dans ce swap à nominal variable. La swaption receveuse de taux fixe, donne le droit à l'acheteur de rentrer dans ce type de swap en payant la jambe variable de flux $\delta L(t, t + \delta) N_t$ et en recevant en contrepartie la série de flux $\delta k N_t$ au taux fixe k fixé en date 0. En date T , il aurait eu la possibilité de rentrer gratuitement dans un swap où il paye cette même jambe variable et reçoit en contrepartie une jambe fixe de flux $SV_T(T_0, T_N, \delta, N_t) \delta N_t$. Par conséquent, cette swaption receveuse de jambe fixe peut aussi être vue comme une option de vente sur le taux de swap forward sur nominal variable $SV_T(T_0, T_N, \delta, N_t)$ de prix d'exercice k . Ainsi, cette swaption sur swap à nominal variable peut s'évaluer sous la forme suivante :

$$\Pi(T, T_0, T_N, \delta, K, N_t) = B(0, T) \mathbb{E}^{Q_T} \left[(K - SV_T(T_0, T_N, \delta, N_t))^+ \sum_{i=1}^N \delta B(T, T_i) N_{T_i} \right].$$

Ce concept de swap à nominal variable est classique en banque mais reste tout de même assez complexe. Le fait que les nominaux varient chaque année donne à ce produit un caractère exotique et rend son étude intéressante comme nous le verrons dans la suite. En utilisant ces concepts classiques, nous allons décrire le produit de transfert du risque de risque de taux inhérent à un portefeuille de rentes.

8.4.3 Étude de la Longevity Nominal Chooser Swaption

Dans l'hypothèse où l'assureur sait prédire à l'avance l'évolution des assurés, il pourrait rentrer dans une swaption à nominal variable qui lui permet de gérer son risque dans le futur. Dans ce produit, on va prendre en compte le fait que l'assureur ne sait pas estimer avec certitude cette évolution mais qu'il sait a priori que la série de nominaux sera comprise entre deux séries extrêmes de nominaux (estimation par des quantiles). Une banque pourra proposer à une assurance de rentrer dans une swaption sur swap à nominal variable où les

nominaux sont donnés par $N_t = \alpha_T N_t^- + (1 - \alpha_T) N_t^+$ et où les courbes N_t^- et N_t^+ sont calculées par un modèle de longévité. Ces courbes extrêmes de nominaux N_t^- et N_t^+ sont fixées au moment de la mise en place du produit et l'assureur aura un degré de liberté en date T via le choix du paramètre α_T .

Plus les courbes extrêmes sont rapprochées, plus le produit se rapproche d'une swaption classique sur swap à nominal variable. Il est tout à fait possible d'indexer la courbe de flux futurs comme barycentre de deux courbes déterminé par un α_T fixé en T . Cela va permettre à l'assureur de gérer le risque de taux inhérent au portefeuille de rentes en tenant compte d'une incertitude sur la longévité des individus et de l'information disponible en date T .

Description du produit

La différence principale entre ce produit et une swaption à nominal variable est que le α_T est déterminé par l'assureur à la date T . La série de nominaux du swap n'est donc connue qu'à la maturité de la swaption dépendant du paramètre α_T . Ainsi, le taux de swap variable dépendra des deux séries extrêmes de nominaux N_t^- et N_t^+ connues et fixées en date 0 ainsi que du paramètre α_T qui lui est calculé en date T . Dans ce cas, on peut réécrire le taux de swap Forward sous la forme suivante :

$$\begin{aligned} SV_T(T_0, T_N, \delta, N_t^-, N_t^+, \alpha_T) = & \alpha_T \frac{\sum_{i=1}^N (N_{T_i}^- - N_{T_{i-1}}^-) (B(T, T_{i-1}) - B(T, T_N))}{\delta \sum_{i=1}^N B(T, T_i) (\alpha_T N_{T_i}^- + (1 - \alpha_T) N_{T_i}^+)} \\ & + (1 - \alpha_T) \frac{\sum_{i=1}^N (N_{T_i}^+ - N_{T_{i-1}}^+) (B(T, T_{i-1}) - B(T, T_N))}{\delta \sum_{i=1}^N B(T, T_i) (\alpha_T N_{T_i}^- + (1 - \alpha_T) N_{T_i}^+)}. \end{aligned}$$

L'introduction de ce facteur α_T choisi par l'assureur en date T permet à l'assureur de tenir compte de l'évolution des individus à partir de la date T sachant une information comme le nombre de survivants à la date T . Pour que l'assureur puisse tenir compte de cela en date 0, on définit une swaption qui donne la possibilité à l'assureur de rentrer dans ce type de swap pour n'importe quel $\alpha \in [0, 1]$ qu'il choisira a priori à la date T . En pratique, l'assureur aura le choix de rentrer dans $n + 1$ séries de nominaux correspondant à $n + 1$ valeurs de α appartenant à l'ensemble $\{\frac{l}{n}, l \in \{0, \dots, n\}\}$ comprises entre 0 et 1. En fait, cela revient à dire qu'en date T , il pourra choisir n'importe quelle série de nominaux du swap comme barycentre des deux séries extrêmes et aura la possibilité de rentrer dans un swap avec ce nominal variable au prix d'exercice K fixé en date 0. Ainsi, on appelle ce produit une *Longevity Nominal Chooser Swaption* dont l'évaluation est donnée par :

$$\frac{\Pi(T, T_0, T_N, \delta, N_t^-, N_t^+, K)}{B(0, T)\delta} = \mathbb{E}^{\mathcal{Q}_T} \left[\max_{0 \leq \alpha_T \leq 1} \left\{ \sum_{i=1}^N B(T, T_i) (\alpha_T N_{T_i}^- + (1 - \alpha_T) N_{T_i}^+) \right. \right. \\ \left. \left. \times (K - SV_T(T_0, T_N, \delta, N_t^-, N_t^+, \alpha_T))^+ \right\} \right].$$

En effet, on prend l'exemple de la compagnie d'assurance qui s'engage à payer la série de rentes $\delta k N_t$ et qui veut se couvrir contre une baisse des taux d'intérêt. Elle pourra donc acheter l'option de rentrer dans un swap à nominal variable qui lui permettra de recevoir une jambe fixe qui délivre une série de flux de rentes sous la forme $\delta k N_t$ à chaque date t alors qu'elle va payer une jambe variable dont les flux sont donnés par $\delta L(t, t + \delta) N_t$ à la date t . Cependant, il faut que le taux de rente k soit du même ordre de grandeur que le taux de swap ce qui n'est pas forcément le cas.

Adaptation du produit aux besoins de l'assureur

Afin que le produit soit adapté aux besoins spécifiques de l'assureur, il faut préciser certains paramètres du produit.

Choix du prix d'exercice K Analysons maintenant le choix du prix d'exercice K permettant à l'assureur de gérer le risque de taux de son portefeuille de rentes. Il y a en effet un problème concernant le choix du prix d'exercice K de la swaption car l'assureur sera intéressé par un prix d'exercice égale au taux de rente du portefeuille mais il faut que le prix d'exercice soit aussi de l'ordre de grandeur du taux de swap pour que le produit ait un sens financier et commercial. En effet, on a considéré implicitement dans ce swap que les taux de rentes et le taux de swap étaient du même ordre de grandeur ce qui n'est pas forcément le cas en pratique. En effet, il est possible que les taux d'intérêt en date 0 soient par exemple observés par le taux de swap ou le taux Libor $L(0, \delta) \sim 2\%$ alors que le taux de rentes est fixé à une valeur sensiblement différente $k = 4\%$. Si on veut mettre en place une swaption sur taux d'intérêt, il faut que le taux de rente soit renormalisé pour être du même ordre de grandeur que le taux de swap.

Facteur de renormalisation Abordons le problème de renormalisation qui est essentiel afin que le produit soit adapté aux besoins de l'assureur. L'assureur souhaiterait idéalement choisir la valeur du prix d'exercice K du produit égale au taux de rentes k , ce qui lui permet de couvrir son risque de taux en rentrant dans une swaption au taux k en recevant la série de flux à taux fixe qu'elle s'est engagée à payer. Le problème est que le taux k peut être très différent du niveau des taux d'intérêt et qu'il ne sera pas possible d'échanger une jambe fixe avec ce taux de rente k contre une jambe variable au taux Libor. Par conséquent, on veut que le taux de rente k soit du même ordre de grandeur que le taux de swap

$SV_0(T_0, T_N, N_t^-, N_t^+, \alpha)$ de sorte que la swaption soit proche de la monnaie et ait une vraie signification financière. Le taux de swap dépend du paramètre α qui n'est pas forcément connu en date 0 mais en pratique, on considère la série de nominaux médiane dont le taux de swap associé est $SV_0(T_0, T_N, N_t^-, N_t^+, 0.5)$. Cette approximation vient du fait que les taux de swap sont généralement assez proches les uns des autres pour toutes les valeurs de α . On définit un ratio de renormalisation $\eta(k)$ par :

$$\eta(k) = \frac{k}{SV_0(T_0, T_N, N_t^-, N_t^+, 0.5)}.$$

Une série de flux de $k\delta N_t$ à la date t peut être représenté par $\eta(k)$ séries identiques de flux $\frac{k}{\eta(k)}\delta N_t = SV_0(T_0, T_N, N_t^-, N_t^+, 0.5)\delta N_t$ à la date t . Ce sont ces nouvelles séries ayant le bon ordre de grandeur que l'on va "swapper" contre des taux Libor. La linéarité des flux permet de réécrire l'évaluation du produit avec ce facteur de renormalisation $\eta(k)$:

$$\begin{aligned} \frac{\Pi(T, T_0, T_N, \delta, N_t^-, N_t^+, k)}{B(0, T)\eta(k)\delta} &= \mathbb{E}^{Q_T} \left[\max_{0 \leq \alpha_T \leq 1} \left\{ \sum_{i=1}^N B(T, T_i) \left(\alpha_T N_{T_i}^- + (1 - \alpha_T) N_{T_i}^+ \right) \right. \right. \\ &\quad \left. \left. \times \left(SV_0(T_0, T_N, \delta, N_t^-, N_t^+, 0.5) - SV_T(T_0, T_N, \delta, N_t^-, N_t^+, \alpha_T) \right)^+ \right\} \right]. \end{aligned}$$

Ce produit sur taux de rente k peut être vu comme $\eta(k) = \frac{k}{SV_0(T_0, T_N, \delta, N_t^-, N_t^+, 0.5)}$ Longevity Nominal Chooser Swaptions au prix d'exercice $SV_0(T_0, T_N, \delta, N_t^-, N_t^+, 0.5)$ et permet de gérer le risque de taux d'intérêt relativement au niveau du taux de rentes k . La renormalisation est donc très facile à mettre en place dans la mesure où le facteur $\eta(k)$ intervient comme facteur multiplicatif du prix de l'option à la monnaie.

Il est possible de choisir un autre facteur de renormalisation, mais celui ci a l'avantage de fournir des swaptions quasiment à la monnaie qui seront par conséquent assez liquides dans le marché. On peut enfin penser à une amélioration qui consiste à regarder en date 0 l'ensemble des $SV_0(T_0, T_N, \delta, N_t^-, N_t^+, \alpha)$ pour chaque valeur de α . À la maturité de l'option, lorsque l'assureur choisira son α , le prix d'exercice de la swaption sera fixé à $SV_0(T_0, T_N, \delta, N_t^-, N_t^+, \alpha)$ ce qui correspond une swaption à la monnaie.

En pratique, l'assureur pourra payer des rentes à taux variables au lieu de rentes à taux fixe si les taux d'intérêt ont baissé. Si les taux d'intérêt augmentent, il n'exercera pas l'option, paiera les rentes dont le taux est "relativement" (à la renormalisation près) plus faible que le niveau des taux d'intérêt et aura a priori un rendement annuel suffisant pour payer les rentes. Ainsi, en achetant ce produit, il se sera couvert contre une baisse des taux d'intérêt. Par le même type argument, l'assureur peut décider de se couvrir contre une hausse des taux d'intérêt (si l'assureur a peur de devoir réajuster les taux de rentes avec une montée de taux d'intérêt) et achètera dans une swaption payeuse de jambe fixe.

8.4.4 Étude qualitative

Ce produit permet à l'assureur de gérer le risque de taux futur lié à son portefeuille de rentes. Bien que ce produit nécessite un "design" particulier du fait de la présence de deux courbes de nominaux spécifiques au contrat, il est plus adapté aux problématiques réelles de l'assureur. En début de contrat, l'assureur fournit à la banque deux courbes de nominaux qu'il estime comme extrêmes afin de mettre en place ce contrat. La banque n'aura pas besoin d'effectuer de calculs supplémentaires dans la mesure où ces courbes de nominaux sont des données d'entrée du produit qui ne modifient pas la méthode d'évaluation. La Longevity Nominal Chooser Swaption est un produit pur de taux d'intérêt qui permet de transférer réellement le risque de taux lié au risque de longévité. La banque n'a pas besoin de gérer un risque quelconque de longévité ou même d'avoir sa modélisation propre de longévité.

En ce qui concerne la maturité T_N , elle doit, en théorie, refléter le moment à partir duquel l'ensemble des individus du portefeuille sont décédés. Cependant, cette maturité est généralement très longue et l'assureur peut considérer une maturité T_n plus faible. Il peut prendre la maturité à partir de laquelle la courbe extrême passe en dessous d'un seuil ϵ d'assurés vivants (10% par exemple) ce qui permettra d'avoir une maturité de swap raisonnable. Les flux de rentes devenant de plus en plus petits, il ne sert à rien de prendre une maturité supérieure à 30 ans pour gérer la longévité de la majeure partie des assurés. Pour des personnes qui auront plus de 70 ans dans 10 ans, une maturité de 30 ans est largement suffisante. Pour des personnes qui auront 60 ans dans 10 ans, on pourra arrêter le produit lorsque les assurés auront un âge proche de 100 ans et il faudra éventuellement prendre une maturité de l'ordre de 40 ans. Dans ce cas, il peut être préférable de prendre une swaption dont la maturité de l'option sera plutôt 15 ans au lieu de 10 ans afin d'avoir une maturité de swap plus courte. Concernant les produits de taux, les banques sont habituées à gérer des maturités de l'ordre de 30 ans alors que les produits d'horizon plus lointains sont généralement trop longs.

D'ailleurs, il faut distinguer les produits dont le paiement contingent se fait par compensation par espèce *cash settlement* des produits où le paiement se fait par règlement physique *physical settlement*. Lorsque le paiement se fait sous le régime de compensation par espèces, on analyse les conditions de marchés et le règlement se fait à la maturité de l'option et le contrat prend fin. Dans le cas d'un paiement par règlement physique, les deux contractants s'échangent les flux jusque la fin de vie du produit. Lorsque les produits ont des maturités longues, on privilégie les contrats dont le paiement se fait par compensation d'espèce afin de réduire l'horizon d'un produit et de limiter les risques de contreparties. Il est donc possible de considérer des swaptions de maturités supérieures à 10 ans avec un tenor de l'ordre de 30 ans à condition que le paiement soit de type "cash settlement".

Concernant l'évaluation de ce produit, il faut utiliser un modèle de taux bien adapté. Étant

donné la structure du produit, un modèle de taux multifactoriel va a priori être plus adapté dans la mesure où ce produit va faire intervenir la corrélation d'un ensemble de zéro-coupon de maturités différentes. Le modèle doit être assez riche pour décrire au mieux ce produit.

Tout d'abord, on se place du point de vue de l'assureur qui cherche à fournir les deux courbes extrêmes concernant la survie de ses assurés. On utilise les modèles microscopiques présentés dans la thèse qui tiennent compte de l'âge des assurés mais aussi d'éventuellement d'autres caractéristiques. On calibre le modèle de mortalité sur les tables françaises de mortalité. Le modèle de dynamique de population fournit un ensemble scénarios d'évolution du nombre d'assurés du portefeuille. On estime la médiane ainsi que différents quantiles de ces scénarios d'évolution représentant le scénario central et les scénarios extrêmes de mortalité (avec des quantiles à 5% et 95%). Les deux courbes de nominaux que l'assureur fournira à la banque seront a priori les quantiles de scénario d'évolution.

Les swaps à nominal variable sont des produits exotiques qui peuvent être vus comme des paniers de swaps. En effet, un swap sur nominal variable (N_t) de maturité T_N revient à considérer un swap de nominal N_{T_N} et de maturité T_N , un swap de nominal $N_{T_{N-1}} - N_{T_N}$ et de maturité T_{N-1}, \dots . Ainsi, le swap à nominal variable est un panier de swaps composé de N swaps de nominaux $N_{T_{N-k}} - \sum_{l=0}^{k-1} N_{T_{N-l}}$ et de maturité T_{N-k} , pour toute valeur de k compris entre 1 et N .

Ainsi, il n'y a pas de prix de marchés pour les swaps à nominal variable et le fait de calibrer le modèle sur certaines swaptions liquides à nominal fixe ne permettra pas de fixer le prix d'un swap à nominal variable. Il sera donc normal que le prix des swaptions sur swap à nominal variable change lorsqu'on fait varier la corrélation entre les taux forward successifs étant donné que ces swaptions sur swap à nominal variable dépendent très fortement de la corrélation entre les taux.

Remarque 8.4. *Afin de gérer la corrélation entre les taux, on pourrait penser à utiliser comme instruments de couverture (donc de calibration) les CMS spreads, mais il apparaît sûrement plus pertinent de traiter la corrélation comme un paramètre "exotique" et de faire des tests pour tester son influence. En effet, les CMS spreads n'étant pas très liquides pour l'ensemble des maturités (le 2Y/10Y est liquide et le 10Y/30Y mais les autres le sont beaucoup moins), il n'est pas forcément pertinent de se calibrer dessus.*

On essaye d'analyser le caractère exotique de la Longevity Nominal Chooser Swaption qui est une option sur maximum. On étudie le comportement du prix du produit en fonction de la corrélation des taux successifs ainsi que les prix des swaptions sur swap variable pour différentes valeurs de α ($\alpha = 0$, $\alpha = 1$ et $\alpha = \alpha_{moy}$).

8.4.5 Étude quantitative

L'euro est la monnaie utilisée par les assureurs européens dont les clients résident en Europe. La courbe des taux européenne a connu des configurations très différentes. On considère par exemple la courbe de taux européenne au 8 Avril 2009 que l'on décrit sur la Figure 8.17.

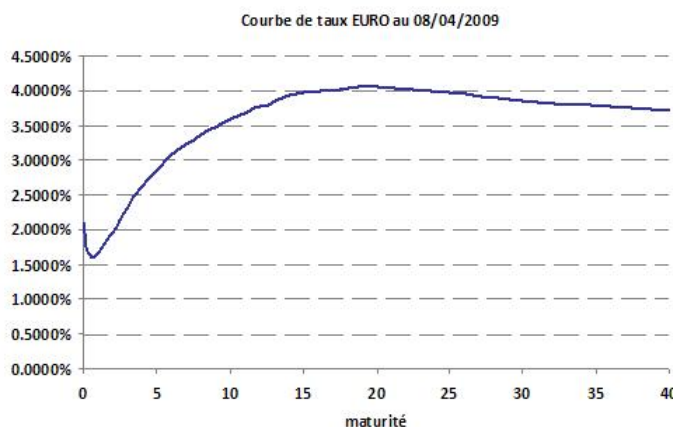


FIGURE 8.17 – Courbe de taux européenne du 8 Avril 2009

Cette courbe de taux révèle une différence importante entre le taux court et le taux long avec des taux longs assez élevés de l'ordre de 4%. Des tests sont effectués pour d'autres configurations de courbes de taux avec des courbes inversées, des courbes contenant des "bosses". Cette configuration fournit des prix assez élevés pour le produit puisque la Longevity Nominal Chooser Swaption permet une couverture contre une baisse de taux d'intérêt avec des taux élevés.

On pourra éventuellement rajouter une autre monnaie et un taux de change dans le cas où une assurance assure des personnes résidant dans un pays où la monnaie est différente. On calibre le modèle de Heath-Jarrow-Morton (HJM) [20] à deux facteurs sur cette courbe de taux ainsi que sur un ensemble de prix de swaptions dont les maturités sont pertinentes dans l'étude du produit. Pour l'évaluation du produit, on considère que le paramètre α entre 0 et 1 appartient à l'ensemble $\{\frac{l}{10}, l \in [0, 10]\}$.

Modèle HJM à deux facteurs Le modèle HJM à deux facteurs associé à une monnaie permet de décrire la dynamique de la courbe des taux dans le temps. Bien que le modèle soit assez flexible pour tenir compte de la corrélation de taux forward, il permet de générer uniquement des smiles normaux pour les options vanilles. Ce modèle est donc assez stan-

dard et simple et décrit une dynamique des taux assez riche en captant de l'information sur la corrélation des taux. On se propose de détailler le modèle un peu plus en détails.

Dans le cadre d'un modèle HJM n facteurs, la dynamique des taux courts forward instantanés est donnée sous la forme :

$$df(t, T) = \mu(t, T)dt + \sigma(t, T)dW_t^\top,$$

où W est un mouvement brownien vectoriel de dimension n sous la probabilité risque neutre \mathbb{Q} , où $\mu(t, T)$ est une fonction réelle représentant le drift du taux court forward et où $\sigma(t, T)$ est un vecteur de taille n qui décrit la volatilité du taux court forward.

Dans cette étude, on a choisi de spécifier une structure de volatilité $\sigma(t, T)$ séparable en dimension 2 sous la forme :

$$\sigma(t, T) = a(t)b(T) \begin{pmatrix} \cos(\theta(t)) \\ \sin(\theta(t)) \end{pmatrix},$$

où $a(t)$, $b(t)$ et $\theta(t)$ sont des fonctions réelles déterministes. Les fonctions a et b sont supposées constantes par morceaux et la fonction θ est supposée linéaire par morceaux. On remarque que la fonction a va caractériser le niveau de volatilité alors que la fonction b va renseigner sur la structure par terme de la volatilité forward. Enfin, la fonction θ permet de capter de l'information sur la corrélation des taux. En effet, on définit la corrélation ρ_i entre deux taux forward successifs et on peut l'exprimer comme suit :

$$\begin{aligned} \rho_i &= \langle f(., T_i), f(., T_{i+1}) \rangle \\ &= \cos(\theta(T_{i+1}) - \theta(T_i)). \end{aligned}$$

On étudie les expressions du prix des zéro-coupon pour en déduire une expression du taux court r_t . On allège les expressions en définissant les notations suivantes :

$$\begin{aligned} \beta_1(t) &= \int_0^t b(s) \cos(\theta(s))ds, & \beta_2(t) &= \int_0^t b(s) \sin(\theta(s))ds, & \alpha(t) &= \int_0^t a^2(s)ds, \\ \lambda_1(t) &= \int_0^t a^2(s)\beta_1(s)ds, & \lambda_2(t) &= \int_0^t a^2(s)\beta_2(s)ds. \end{aligned}$$

Partant de cette spécification du modèle HJM, on peut en déduire la dynamique de la courbe de prix des zéro-coupon sous la probabilité risque neutre \mathbb{Q} est donnée par la diffusion suivante :

$$\frac{dB(t, T)}{B(t, T)} = r_t dt - a_t (\beta_1(T) - \beta_1(t)) dW_1^\mathbb{Q}(t) - a(t) (\beta_2(T) - \beta_2(t)) dW_2^\mathbb{Q}(t),$$

où $W_1^{\mathbb{Q}}$ et $W_2^{\mathbb{Q}}$ sont des mouvements browniens indépendants. La dynamique du prix des zéro-coupon incite à définir de manière naturelle deux facteurs qui pilotent les courbes de prix des zéro-coupon :

$$X_t^1 = \int_0^t a(s) dW_1^{\mathbb{Q}}(s), \quad X_t^2 = \int_0^t a(s) dW_2^{\mathbb{Q}}(s).$$

On en déduit une forme explicite du prix des zéro-coupon en fonction des deux facteurs X^1 et X^2 du modèle :

$$B(t, T) = \frac{B(0, T)}{B(0, t)} e^{-(\beta_1(T) - \beta_1(t))[\frac{1}{2}\beta_1(T) + \beta_1(t)\alpha(t) - \lambda_1(t) + X_t^1] - (\beta_2(T) - \beta_2(t))[\frac{1}{2}\beta_2(T) + \beta_2(t)\alpha(t) - \lambda_2(t) + X_t^2]}.$$

Par reconstruction, le taux court peut aussi s'exprimer en fonction des facteurs :

$$r_t = f(0, t) + b(t) \cos(\theta(t))[\alpha(t)\beta_1(t) - \lambda_1(t) + X_t^1] + b(t) \sin(\theta(t))[\alpha(t)\beta_2(t) - \lambda_2(t) + X_t^2],$$

où $f(0, t)$ représente le taux court forward instantané. Le taux court apparaît donc comme une forme linéaire des facteurs du modèle.

Du point de vue de la calibration, on calibre tout d'abord la fonction θ qui pilote la corrélation des taux. On verra que la corrélation des taux va jouer un rôle central dans l'étude de la Longevity Nominal Chooser Swaption. Après avoir calibré la courbe θ , on calibre les courbes a et b sur un ensemble d'options vanilles. On choisira généralement les options vanilles qui seront les mieux adaptées au produit exotique étudié.

Courbes de nominaux extrêmes On associe le modèle de mortalité au modèle de dynamique de population afin de générer les scénarios d'évolution. On donne pour information la série médiane de nominaux ainsi que les séries extrêmes de nominaux pour 62482 hommes français en 2009 calculées grâce au modèle de dynamique de population. Les quantiles à 5%, 50% et 95% sont calculés avec 10000 scénarios d'évolution (voir Annexe B).

Choix du prix d'exercice Il est difficile de fixer le prix d'exercice K de sorte que la swaption choisie soit à la monnaie. En effet, le choix du paramètre α se fait en date T ce qui modifie la valeur du taux de swap. Cependant, on considère un prix d'exercice $K = SV_0(T_0, T_N, \delta, N_t^-, N_t^+, \alpha) = 4.3\%$ qui sera en pratique le niveau à la monnaie des swaptions quelque soit la valeur du paramètre α . En effet, avec les deux courbes extrêmes précédentes, on peut constituer deux swaps à nominal variable dont le premier est associé à la série inférieure de nominaux et le deuxième associé à la série supérieure de nominaux. Le taux de swap associé à la série de nominal inférieure est de l'ordre de 4.35% alors que celui associé à la série de nominal supérieur est de l'ordre de 4.25%. On peut voir que le

taux $k = 4.3\%$ permet d'égaliser la jambe fixe et la jambes variable pour ces deux swaps extrêmes à nominal variable :

$$\begin{aligned} \text{Courbe inférieure}(\alpha = 1) : & \quad BF^{inf} = 13055 \text{ BP} , & \quad BV^{inf} = 13394 \text{ BP} . \\ \text{Courbe supérieure}(\alpha = 0) : & \quad BF^{sup} = 16750 \text{ BP} , & \quad BV^{sup} = 16534 \text{ BP} . \\ \text{Courbe centrale}(\alpha = 0.5) : & \quad BF^{mid} = 14915 \text{ BP} , & \quad BV^{mid} = 15008 \text{ BP} . \end{aligned}$$

Notion de "coût sur la rente" L'évaluation de la Longevity Nominal Chooser Swaption permet de se rendre compte du prix de ce produit de couverture. On introduit la notion de "coût sur la rente" afin de rendre plus intuitif la lecture des prix. En effet, si on considère un taux de rentes à $k\%$, le prix de la rente sur une série de nominal indexé par α

$$P_{rente} = k * LVL(\alpha),$$

où α décrit la série de nominal et où $LVL(\alpha) = \sum_{i=1}^N (\alpha N_{T_i}^- + (1-\alpha) N_{T_i}^+) B(T, T_i)$ correspond au niveau d'actualisation des flux. On définit la notion de coût en fonction du prix du produit avec une actualisation avec un niveau central de α

$$P_{produit} = c * LVL(\alpha = 0.5).$$

Ainsi, la valeur de c permet d'estimer le surcoût lié à l'achat du produit dans le cas où l'assureur n'a pas besoin d'exercer l'option. En effet, il aura payé un taux de rente effectif de $k + c$ au lieu d'un taux de rente de k . Cette notion apparaît être une notion naturelle afin d'estimer l'ordre de grandeur de la prime par rapport au niveau de rente.

Étant donné la méthode de renormalisation, on peut supposer que le taux de rente est égale à $k = 4.3\%$ et le calcul du coût c sera à mettre en regard avec cette valeur de taux de rente.

Impact de la corrélation des taux

Ce produit est un produit sur maximum qui possède des propriétés intéressantes. Tout d'abord, on sait que le prix de ce produit est naturellement plus élevé que la plus chère des swaptions à nominal variable car l'inégalité de Jensen appliquée au maximum fournit $\mathbb{E}(\max(A)) \geq \max(\mathbb{E}(A))$. De plus, ce produit va dépendre de la corrélation qu'il existe entre différents taux de swaps à des maturités différentes.

Dans le modèle de HJM 2 facteurs, il est possible de caractériser la corrélation successive des taux spot forward successifs. Au Crédit Agricole CIB, la spécification utilisée est

$$\begin{aligned} \rho &= \text{correl}(f(., T), f(., T + dt)) \\ &= \rho_\infty + (\rho_\infty - \rho_0) \exp(-0.20dt) \\ &= \cos(\theta(T + dt) - \theta(T)). \end{aligned}$$

L'idée est qu'en jouant sur ρ , on peut modifier la corrélation entre les taux spot forward pour différentes maturités. En effet, le fait de prendre un produit qui tient compte d'un maximum donne l'intuition que plus ces taux sont décorrélés, plus le prix du produit va être grand. D'un point de vue pratique, on posera $\rho_0 = \rho_\infty$ et on fera varier le paramètre $\rho = \rho_0$.

Il faudra aussi garder en tête qu'on souhaite que le taux court et le taux long soient les plus décorrélés possibles (ce qui est assez intuitif dans la mesure où ce qu'il va se passer dans 45 ans ne doit pas dépendre énormément des taux actuels. Cette corrélation peut être analysée au travers de l'angle θ et il faudra donc que la courbe de θ à l'infini soit pas trop éloignée de l'angle π ce qui correspond à une décorrélation parfaite.

Afin de trouver les taux de swap qui interviennent dans ce produit, on utilise les courbes de nominaux. On cherche à voir quelle était la maturité moyenne des swaps à nominal variable concernant les deux swaps à nominal variable (issus des courbes de nominaux extrêmes). Ces maturités moyennes peuvent être calculées sous la forme :

$$T_{moy}^N = \frac{\sum_{i=1}^N T_i N_{T_i}}{\sum_{i=1}^N N_{T_i}}.$$

En considérant les deux séries de nominaux, on a trouvé des swaps de maturités $T_{moy}^{N-} = 10Y$ pour la courbe inférieures et $T_{moy}^{N+} = 12Y$ pour la courbe supérieure. En jouant sur la corrélation des taux comme précédemment, on va modifier la corrélation entre le taux de swap de maturités 10Y et le taux de swap de maturité 12Y. On peut comment varie la corrélation entre les deux swaps possédant ces deux maturités et voir si elle a un impact. On analyse le produit dans le tableau suivant (en faisant varier ρ) de telle sorte à conserver une courbe d'angle θ raisonnable.

Corrélation taux successifs	Corrélation swap 10Y/12Y	Prix $\alpha = 0$	Prix $\alpha = 0.5$	Prix $\alpha = 1$	Prix du produit	Coût sur rente
$\rho_0 = 0.998$	99.6%	3015 bp	2656 bp	2244 bp	3020 bp	0.87%
$\rho_0 = 0.997$	99.3%	2954 bp	2580 bp	2215 bp	2960 bp	0.853%
$\rho_0 = 0.996$	99.1%	2886 bp	2529 bp	2183 bp	2897 bp	0.835%
$\rho_0 = 0.995$	98.8%	2813 bp	2474 bp	2147 bp	2828 bp	0.815%
$\rho_0 = 0.994$	98.6%	2732 bp	2412 bp	2107 bp	2751 bp	0.793%
$\rho_0 = 0.993$	98.3%	2641 bp	2342 bp	2061 bp	2667 bp	0.769%
$\rho_0 = 0.992$	98%	2540 bp	2264 bp	2009 bp	2574 bp	0.742%

TABLE 8.1 – Évolution du prix du produit en fonction de la corrélation des taux successifs

Ces différentes valeurs permettent de décorréler de plus en plus les taux courts successifs et on voit que plus les taux sont décorrélés plus le prix du produit augmente. Cela s'explique

par le fait qu'on considère un produit sur maximum et que lorsque la corrélation est petite, il y a moins de contraintes et il peut se passer plus de phénomènes en autorisant des scénarios plus "variés". Il ne faut pas oublier qu'on ne peut pas trop décorréler les taux dans la mesure où il faut avoir une courbe θ raisonnable de manière à ce que les taux longs soient largement décorrélés des taux courts en ayant par exemple un $\theta_{45} - \theta_0 = \theta_{45}$ de l'ordre de π ou $\frac{3\pi}{2}$.

On se rend compte que la corrélation va jouer un rôle très important pour l'évaluation du produit. Elle va à la fois jouer sur le prix des marginales (swaptions à α fixé) mais va avoir aussi un impact sur le prix du produit. En effet, plus la corrélation diminue plus le prix des marginales diminue et plus le prix du produit diminue. On pourrait penser que cela est contradictoire avec l'impact intuitif que plus la corrélation diminue, plus le prix d'un produit sur maximum augmente. En fait, ici le prix des différentes swaptions à nominal variable varie aussi avec la corrélation et il ne faut pas raisonner en terme de prix mais en différence de prix entre le produit et les marginales. Par exemple, on peut introduire la notion de *switch option* correspondant à la différence entre le prix du produit et la plus chère des marginales. Le prix de la *switch option* permet de quantifier le caractère exotique du produit.

On peut voir que lorsque la corrélation diminue, le prix de chacune des swaptions diminue et l'écart de prix entre les différentes swaptions diminue aussi. Cependant, on observe que l'écart entre le prix du produit et la plus chère des swaptions augmente fortement bien que le prix des swaptions diminue. En effet, on voit qu'à la corrélation $\rho_0 = 0.998$, l'écart de prix entre les swaptions sur nominaux extrêmes ($\alpha = 0$ et $\alpha = 1$) est de $3015 - 2244 = 771$ BP et que le prix de la *switch option* est de $3020 - 3015 = 5BP$. On voit que lorsque la corrélation diminue, l'écart entre les prix de swaptions diminue alors que l'écart entre le prix du produit et le prix de la swaption la plus chère augmente. À la corrélation $\rho_0 = 0.992$, la différence de prix entre les swaptions sur nominaux extrêmes est de $2540 - 2009 = 531$ BP alors que le prix de la *switch option* $2574 - 2540 = 34$ BP. Cela confirme le fait que la corrélation a un impact sur le caractère exotique du produit et que plus la corrélation diminue, plus le produit a un comportement exotique.

Ainsi, bien que le prix des différentes swaptions diminue avec la corrélation, on voit que le prix du produit devient de plus en plus cher relativement au prix des swaptions. C'est ce phénomène qu'il faut observer et qui ne contredit plus l'impact intuitif de la corrélation qui suggère que la valeur d'un produit sur maximum augmente lorsque la corrélation diminue. Pour décrire le plus précisément possible ce produit, il faudra un modèle qui décrive au mieux les corrélations entre les taux de swap pour des maturités différentes et l'adaptation d'un modèle affine comme celui qui est détaillé dans le chapitre 3 pourrait être une bonne solution.

On observe que la différence de prix n'est pas très grande sur ce produit et qu'elle permet à l'assureur de choisir le α qui lui conviendra précisément. Pour des valeurs extrêmes de corrélation, le prix du produit ne vaut que $\frac{2574-2540}{2540} = 1.3\%$ plus cher que la plus chère des swaptions ce qui tend à laisser penser que le produit pourrait intéresser un assureur. On peut aussi noter le fait que le prix du produit sera d'autant moins cher que les courbes de nominaux ne seront pas trop extrêmes. Dans le cas limite où les deux courbes extrêmes sont confondues en une seule série de nominaux, on vérifie facilement que le prix du produit vaut le prix de l'unique swaption sur nominal variable.

Cependant, on peut remarquer que la swaption la plus chère est toujours la swaption pour $\alpha = 0$ qui correspond à la swaption sur swap avec la série supérieure de nominaux. Le fait de changer la corrélation ne change pas l'ordre des prix et la swaption pour $\alpha = 0$ reste la plus chère quelque soit la corrélation des taux. On va essayer de voir si le prix d'exercice peut changer l'ordre des prix de swaption et quel est son impact sur le produit.

Impact du prix d'exercice

Bien que le produit soit a priori destiné à être utilisé à la monnaie étant donné une valeur du taux de rente du portefeuille, on peut aussi regarder le comportement du produit en fonction du prix d'exercice. On fixe la corrélation à $\rho_0 = 0.995$ et on analyse l'impact du prix d'exercice en considérant une large plage de valeurs.

Corrélation taux successifs	Prix d'exercice	Prix $\alpha = 0$	Prix $\alpha = 0.5$	Prix $\alpha = 1$	Prix du produit	Coût sur rente
$\rho_0 = 0.995$	$K = 1\%$	51 bp	56 bp	63 bp	68 bp	0.01%
$\rho_0 = 0.995$	$K = 2\%$	257 bp	250 bp	248 bp	285 bp	0.08%
$\rho_0 = 0.995$	$K = 3\%$	882 bp	807 bp	738 bp	910 bp	0.26%
$\rho_0 = 0.995$	$K = 4\%$	2236 bp	1977 bp	1730 bp	2254 bp	0.65%
$\rho_0 = 0.995$	$K = 5\%$	4466 bp	3895 bp	3341 bp	4471 bp	1.28%
$\rho_0 = 0.995$	$K = 6\%$	7458 bp	6485 bp	5529 bp	7461 bp	3.16%
$\rho_0 = 0.995$	$K = 7\%$	10949 bp	9537 bp	8137 bp	10949 bp	5.15%

TABLE 8.2 – Évolution du prix du produit en fonction du prix d'exercice

Ce n'est pas forcément la swaption sur la série supérieure de nominaux qui donne le prix le plus élevé puisqu'on observe une inversion d'ordre de prix pour un prix d'exercice compris entre $K = 1\%$ et $K = 2\%$. En effet, pour les prix d'exercices supérieurs à 1.8%, la swaption sur le swap avec la série supérieure de nominaux est la plus chère. En dessous de cette valeur, c'est la swaption sur le swap avec la série inférieure de nominaux qui devient la plus chère. On représente cette évolution sur le tableau suivant.

Corrélation taux successifs	Prix d'exercice	Prix $\alpha = 0$	Prix $\alpha = 0.5$	Prix $\alpha = 1$	Prix du produit	Coût sur rente
$\rho_0 = 0.995$	$K = 1\%$	51 bp	56 bp	63 bp	68 bp	0.01%
$\rho_0 = 0.995$	$K = 1.1\%$	61 bp	66 bp	74 bp	79 bp	0.023%
$\rho_0 = 0.995$	$K = 1.2\%$	73 bp	77 bp	85 bp	92 bp	0.026%
$\rho_0 = 0.995$	$K = 1.3\%$	87 bp	91 bp	98 bp	107 bp	0.031%
$\rho_0 = 0.995$	$K = 1.4\%$	103 bp	106 bp	113 bp	124 bp	0.035%
$\rho_0 = 0.995$	$K = 1.5\%$	121 bp	123 bp	130 bp	144 bp	0.0415%
$\rho_0 = 0.995$	$K = 1.6\%$	142 bp	143 bp	149 bp	166 bp	0.046%
$\rho_0 = 0.995$	$K = 1.7\%$	165 bp	166 bp	170 bp	190 bp	0.054%
$\rho_0 = 0.995$	$K = 1.8\%$	192 bp	191 bp	193 bp	218 bp	0.063%
$\rho_0 = 0.995$	$K = 1.9\%$	223 bp	219 bp	220 bp	250 bp	0.072%
$\rho_0 = 0.995$	$K = 2\%$	257 bp	250 bp	248 bp	285 bp	0.082%

TABLE 8.3 – Évolution du prix du produit en fonction du prix d'exercice

On constate que l'ordre est modifié à partir du prix d'exercice de 1.8% qui est assez éloigné de la monnaie à 4.3%. Bien que l'ordre des prix ait changé, on observe que la plus chère et la moins chère des swaptions correspondent aux swaption associées aux courbes extrêmes de nominaux. Ceci est intuitif puisque selon la valeur de $(K - SV_T(T_0, T_N, \delta, \alpha))^+$, il sera plus intéressant de considérer le swap supérieur ou le swap inférieur. En dessous du prix d'exercice 1.8%, le swap le plus cher est le swap inférieur caractérisé par $\alpha = 1$. Lorsque le prix d'exercice atteint la valeur de 1.8%, les prix des deux swaptions sont du même ordre et le swap le plus cher devient le swap supérieur caractérisé par $\alpha = 0$.

On donne une explication à l'existence de ce prix d'exercice "pivot" ainsi que de son ordre de grandeur. Pour cela, on fait une analogie dans le cas de la "Longevity Maturity Chooser Swaption" qui était un produit pour lequel la série de nominaux était fixe. En effet, dans le cas de la Longevity Maturity Chooser Swaption, on pouvait avoir une intuition sur la maturité de swap qui fournira le maximum du pay off.

On rappelle le pay off de la Longevity Maturity Chooser Swaption sous la probabilité forward neutre donnée en Section 3.2.2 :

$$\Phi = \max_{T_{min} \leq T_N \leq T_{max}} \left\{ (K - S_T(T_0, T_N, \delta))^+ \sum_{i=1}^N \delta B(T, T_i) \right\}.$$

Ce pay off peut se réécrire en le voyant comme une option sur zéro-coupon

$$\Phi = \max_{T_{min} \leq T_N \leq T_{max}} \left\{ (K\delta \sum_{i=1}^N B(T, T_i) + B(T, T_N) - B(T, T_0))^+ \right\}.$$

Le terme $O_N = K\delta \sum_{i=1}^N B(T, T_i) + B(T, T_N)$ correspond au prix d'une obligation en T commençant en T_0 et se terminant en T_N et il faut étudier sa monotonie est à étudier en fonction de la maturité T_N . Il apparaît intéressant d'étudier la différence $O_N - O_{N-1}$

$$\begin{aligned} O_N - O_{N-1} &= K\delta B(T, T_N) + B(T, T_N) - B(T, T_{N-1}) \\ &= B(T, T_{N-1})[B_T(T_{N-1}, T_N)(\delta K + 1) - 1]. \end{aligned}$$

Ainsi, le changement de signe de $O_N - O_{N-1}$ est liée à la position de $L_T(T_{N-1}, T_N) = \frac{1}{\delta} \left(\frac{1}{B_T(T_{N-1}, T_N)} - 1 \right)$ par rapport à K . C'est d'ailleurs assez intéressant de voir que pour un produit qui fait intervenir des swaps dont les maturités sont longues, le comportement du produit dépend de la position du prix d'exercice par rapport à un taux COURT forward $L_T(T_{N-1}, T_N)$.

En regardant le prix de l'obligation de manière continue, on a :

$$O_N = \delta K \int_T^{T_N} B(T, s) ds + B(T, T_N).$$

Étant donné $B(T, T_N) = e^{-\int_T^{T_N} f(T, s) ds}$, la maturité donnant l'obligation la plus chère est caractérisée par une annulation de la dérivée $\frac{\partial O_N}{\partial T_N}$ et est donnée par la relation

$$\delta K B(T, T_N) = B(T, T_N) f(T, T_N),$$

correspondant à $f(T, T_N) \sim \delta K$ le taux court forward. Ainsi, le max du produit sera a priori atteint pour la maturité T_N telle que $f(T, T_N) = K$. Ce produit dépend de la dynamique des taux courts forward.

Ainsi, on peut faire une analogie "intuitive" dans le cas du produit sur des swap à nominaux variables. Lorsque le prix d'exercice dépasse une certaine valeur de l'ordre de grandeur d'un taux court forward, la swaption la plus chère passe de la maturité la plus courte à la plus longue et donc de la série inférieure de nominaux à la série supérieure de nominaux. La courbe des taux que l'on a considéré (voir Figure 8.17) était de l'ordre de 1.8% et cela explique l'ordre de grandeur du prix d'exercice à partir duquel l'ordre des options change.

8.4.6 Piste d'amélioration de la modélisation

Nous avons introduit la Longevity Nominal Chooser Swaption qui permet un transfert intéressant du risque de taux de l'assurance vers une banque. Ce produit est décrit en détails mais il est possible de considérer des extensions intéressantes.

Hétérogénéité du portefeuille

Le modèle de dynamique de population permet de prendre en compte l'évolution des individus. Nous avons traité l'hétérogénéité en âge des assurés mais il peut aussi y avoir

une hétérogénéité liée aux caractéristiques des assurés. De plus, un portefeuille peut être inhomogène avec des rentes individuelles différentes que ce soit par leur nominaux ou leur taux de rentes.

Dans le modèle microscopique, il est possible d'associer un nominal à chacun des individus et nous pouvons très facilement généraliser le produit que l'on a étudié au cas de portefeuille inhomogène par nominaux. Au décès de l'individu, l'ensemble des flux de rentes à venir baisse du taux de rente multiplié par le nominal de l'individu décédé.

Cependant, nous avons supposé dans l'étude du produit que le taux de rente était le même pour l'ensemble des individus du portefeuille. Étant donné que c'est la seule supposition que l'on a faite, nous pouvons supposer que les assurances peuvent ordonner leurs portefeuilles de rentes par taux de rente. D'ailleurs, cela est souvent le cas en pratique. Si cela n'est pas possible, il faudra alors estimer un taux de rente k^* moyen dans le cas où les assurés ont des taux de rente assez proches les uns des autres. Il faudrait tester le modèle microscopique et la Longevity Nominal Chooser Swaption sur un portefeuille de marché complet avec l'ensemble de l'information disponible. L'assureur pourrait ainsi gérer le risque de longévité systémique, le risque de base et le risque de taux d'intérêt.

Enfin, on a considéré le cadre d'une population fermée sans tenir compte des nouveaux assurés. La modélisation de l'immigration dans le modèle de dynamique de population (Chapitre 6) permettrait de tenir compte d'un flux de nouveaux assurés dans le temps.

Étude du produit avec un modèle de taux complexe

Nous avons étudié le produit avec un modèle de HJM 2 facteurs qui a donné des résultats sur l'impact de la corrélation des taux forward successifs. Il serait intéressant de tester ce produit avec un modèle plus complexe, comme le modèle de Wishart décrit dans le chapitre 3, afin de prendre en compte efficacement l'enjeu des corrélations entre les taux de swap qui jouent manifestement un rôle important dans l'évaluation du produit.

Cependant, l'évaluation d'un tel produit avec des modèles complexes n'est pas évident. Si nous utilisons le modèle de Wishart, il faudrait trouver une méthode efficace d'évaluation. Une méthode numérique complexe de type Monte Carlo n'est pas forcément la méthode la plus simple à mettre en place avec les processus de Wishart.

Risque de défaut de la banque

Dans la situation actuelle, le risque de crédit est un risque qu'il faut essayer de prendre en compte. En effet, les nominaux qui entrent en jeu peuvent être très élevés et la maturité du produit est importante ce qui augmente le risque de contrepartie.

Nous pourrions appliquer certaines méthodes classiques en risque de crédit qui tiennent compte du rating des entreprises par les agences de notations. Cependant, la crise récente a démontré que ce rating n'était pas fiable et qu'il ne renseignait pas forcément de la solidité de l'entreprise. Il faudrait donc trouver des méthodes afin de tenir compte du risque de défaut des entreprises.

Risque de liquidité

L'impact de la taille des nominaux mis en jeu peut être essentiel et il faut en tenir compte. En effet, nous savons que lorsque les transactions se font sur des tailles importantes, il peut y avoir des problèmes de liquidité liés à la profondeur de marché. D'ailleurs, la crise financière de 2008 est aussi venue du fait que la plupart des banques avaient la même position sur des tailles importantes et ne pouvaient plus couvrir leurs positions.

Le marché de la longévité est un marché sur lequel les nominaux peuvent potentiellement atteindre des montants très élevés et il faudrait mener une étude sur sa profondeur de marché. Le prix que l'on a donné pour la "Life Nominal Chooser Swaption" est donc un prix "minimal" qui ne tient pas compte du risque de liquidité.

8.5 Conclusion

Dans ce chapitre, nous avons étudié les applications financières des modèles décrits dans les chapitres précédents de la thèse. Nous avons vu qu'il était possible de caractériser le risque systémique de longévité et que les modèles microscopiques fournissaient une information importante sur le risque spécifique de longévité.

Nous avons observé que le risque pur de longévité était compliqué à transférer étant donné l'absence d'indice de référence entre deux intervenants. Cependant, nous avons étudié l'éventuel transfert du risque de taux inhérent aux produits d'assurances vie. Nous avons décrit ce risque de taux et avons étudié la gestion de celui-ci. L'étude de certains produits nous a permis de décrire les qualités qu'un produit de transfert de risque devait avoir.

Ainsi, nous avons proposé la Longevity Nominal Chooser Swaption qui est le produit qui nous semble être le plus adapté aux besoins des deux intervenants. Nous avons décrit en détails ce produit et nous avons effectué des tests concernant l'influence de paramètres tels que la corrélation des taux et le prix d'exercice de la swaption. Nous avons utilisé un modèle de taux assez simple (HJM 2 Facteur) afin de bien comprendre les caractéristiques du produit.

Il existe de nombreuses extensions possibles à l'étude de ce produit. L'extension principale

serait de répondre à la problématique du risque de crédit qui est un enjeu majeur dans les produits sur longévité. Il serait aussi intéressant de considérer l'évaluation du produit avec un modèle de taux plus riche afin de décrire au mieux le risque de taux, de volatilité et de corrélation.

8.6 Annexes

Annexe A : Distribution de l'âge des assurés

Âge	Effectif	Âge	Effectif
55	607	78	998
56	1080	79	923
57	1302	80	858
58	1397	81	836
59	1439	82	835
60	3543	83	806
61	3829	84	667
62	3704	85	522
63	3608	86	354
64	3526	87	248
65	3554	88	245
66	3581	89	226
67	3385	90	209
68	3130	91	185
69	2825	92	141
70	2516	93	109
71	2279	94	78
72	2004	95	52
73	1742	96	36
74	1521	97	22
75	1300	98	14
76	1155	99	9
77	1078	100	4

TABLE 8.4 – Répartition des âges du portefeuilles de rentes

Annexe B : Statistiques d'évolution des assurés du portefeuille

Année	Scénario inférieur	Scénario Médian	Scénario supérieur	Année	Scénario inférieur	Scénario Médian	Scénario supérieur
2009	62482	62482	62482	2038	12124	16457	20825
2010	61259	61414	61519	2039	10553	14813	19214
2011	59984	60304	60529	2040	9136	13229	17642
2012	58675	59150	59498	2041	7826	11749	16097
2013	57302	57961	58433	2042	6629	10351	14646
2014	55881	56723	57327	2043	5546	9063	13233
2015	54405	55444	56182	2044	4608	7862	11902
2016	52888	54120	54995	2045	3767	6773	10674
2017	51332	52755	53782	2046	3040	5783	9501
2018	49721	51343	52521	2047	2435	4896	8414
2019	48049	49888	51223	2048	1910	4110	7413
2020	46341	48388	49895	2049	1483	3421	6496
2021	44588	46841	48520	2050	1132	2818	5666
2022	42814	45251	47109	2051	860	2302	4938
2023	40951	43619	45665	2052	636	1865	4258
2024	39076	41942	44170	2053	465	1498	3656
2025	37162	40223	42629	2054	333	1192	3134
2026	35176	38468	41078	2055	235	941	2674
2027	33154	36678	39482	2056	164	737	2261
2028	31164	34860	37850	2057	112	572	1911
2029	29126	33009	36179	2058	76	441	1605
2030	27111	31145	34487	2059	51	337	1350
2031	25062	29262	32820	2060	33	259	1132
2032	23074	27368	31092	2061	21	195	945
2033	21112	25497	29344	2062	14	146	788
2034	19171	23614	27592	2063	8	108	658
2035	17295	21750	25866	2064	5	81	540
2036	15485	19930	24140	2065	3	60	446
2037	13752	18169	22471	2066	0	43	367

TABLE 8.5 – Scénario d'évolution du portefeuille réel en 2009

Bibliographie

- [1] Antolin P. and H. Blommestein (2007), “Governments and the market for longevity-indexed bonds”, *OECD Journal : Financial Market Trends*, 4, 92.
- [2] Barrieu P., H. Bensusan, N. El Karoui, C. Hillairet, S. Loisel, C. Ravanelli and Y. Slahi (2010), “Understanding, Modeling and Managing Longevity Risk : Key Issues and Main Challenges”, *Scandinavian Actuarial Journal*, Accepted
- [3] Baxter S.D., A.T. Gaches, A.M. Madrigal, F.E. Matthews and D.D. Patel (2009), “Chat longevity predictors should be allowed for when valuing pension scheme liabilities”, *Institute of Actuaries and Faculty of Actuaries*, September.
- [4] Brouhns N. and M. Denuit (2002), “Risque de longévité et rentes viagères. Tables de mortalité prospectives pour la population belge”, *Belgian Actuarial Bulletin*, 2, 1.
- [5] Biffis E. (2005), “Affine processes for dynamic mortality and actuarial valuations”, *Insurance : Mathematics and Economics*, 37, 3, 443-468.
- [6] Blake D., A. Cairns and K. Dowd (2006), “Living with mortality : Longevity bonds and other mortality-linked securities”, *British Actuarial Journal*, 12, 153-228.
- [7] Blake D., A. Cairns and K. Dowd (2006), “A two-factor model for stochastic mortality with parameter uncertainty : Theory and calibration”, *Journal of Risk and Insurance*, 73, 687-718
- [8] Blake D., A. Cairns, K. Dowd, G.D. Coughlan, D. Epstein, A. Ong and I. Balevich (2007), “A quantitative comparison of stochastic mortality models using data from England & Wales and the United States”, Working paper, Heriot-Watt University, and Pensions Institute Discussion Paper PI-0701.
- [9] Brigo D. and F. Mercurio (2005), “Interest Rate Models-theory and Practice : With Smile, Inflation and Credit”, Springer.
- [10] Cox S.H and Y. Lin (2008), “Securitization of catastrophe mortality risks”, *Insurance : Mathematics and Economics*, 42, 2, 628-637.
- [11] Cox S.H., Y. Lin, and H. Pedersen (2010), “Mortality risk modeling : Applications to insurance securitization”, *Insurance : Mathematics and Economics*, 46, 1, 242-253.
- [12] Cox S.H. and Y. Lin (2007), “Natural hedging of life and annuity mortality risks”, *North American Actuarial Journal*, 11, 3, 1-15.

- [13] Dahl M. (2004), "Stochastic mortality in life insurance : market reserves and mortality-linked insurance contracts", *Insurance : Mathematics and Economics*, 35, 1, 113-136.
- [14] Dahl M. and T. Møller (2006), "Valuation and hedging of life insurance liabilities with systematic mortality risk", *Insurance : Mathematics and Economics*, 39, 2, 193-217.
- [15] Faucillon L., M. Juillard and F. Planchet (2006), "Quantification du risque systématique de mortalité pour un régime de rentes en cours de service", Working Paper, ISFA.
- [16] Ferrière R. and V.C. Tran (2009), "Stochastic and deterministic models for age-structured populations with genetically variable traits", *ESAIM : Proceedings*, Vol 27, 289-310.
- [17] Fournier N. and S. Méléard (2004), "A microscopic probabilistic description of a locally regulated population and macroscopic approximations", *Annals of Applied Probability*, 14, 4, 1880-1919.
- [18] Gourieroux C. and A. Monfort (2008), "Quadratic stochastic intensity and prospective mortality tables", *Insurance : Mathematics and Economics*, 43, 1, 174-184.
- [19] Håri N., A. De Waegenaere, B. Melenberg and T.E. Nijman (2008), "Longevity risk in portfolios of pension annuities", *Mathematics and Economics*, 42, 2, 505-519.
- [20] Heath D., A. Jarrow and A. Morton (1992), "Bond Pricing and the Term Structure of Interest Rates : A New Methodology for Contingent Claims Valuation", *Econometrica*, 60, 1, 77-105.
- [21] Hodges S.D. and A. Nauberger (1989), "Optimal replication of contingent claims under transaction costs", *Review of Futures Markets*, 8, 2, 222-239.
- [22] Imhoff E.V. and W. Post (1998), "Microsimulation methods for population projection", *Population*, 1, 97-136.
- [23] Jusot F. (2004), "Mortalité, revenu et inégalités de revenu en France : une analyse multi-niveaux", IRDES, Reference BDSP : 309602.
- [24] Lee R.D. and L.R. Carter (1992), "Modeling and forecasting U.S. mortality", *Journal of the American Statistical Association*, 87, 659-675.
- [25] Lee R.D. and T. Miller (2001), "Evaluating the performance of the Lee-Carter model for forecasting mortality", *Demography*, 38, 537-549.
- [26] Menoncin F. (2008), "The role of longevity bonds in optimal portfolios", *Insurance : Mathematics and Economics*, 42, 1, 343-358.
- [27] Milevsky M.A. and Promislow S.D. (2001), "Mortality derivatives and the option to annuitise", *Insurance : Mathematics and Economics*, 29, 3, 299-318.
- [28] Nguyen T.L. (2011), "Risque financier en assurance, variables annuités et risque long terme", en préparation.
- [29] Plat R. (2009), "Stochastic portfolio specific mortality and the quantification of mortality basis risk", *Insurance : Mathematics and Economics*, 45, 1, 123-132.

- [30] Schrager D.F. (2006), “Affine Stochastic mortality”, *Insurance : Mathematics and Economics*, 38, 1, 81-97.
- [31] Tran V.C. (2006), “Modèles particuliers stochastiques pour des problèmes d’évolution adaptative et pour l’approximation de solutions statistiques”, Dissertation.
- [32] Tran V.C. (2008), “Large population limit and time behaviour of a stochastic particle model describing an age-structured population”, *ESAIM : P&S*, Vol 12, 345-386.
- [33] Von Foerster H., (1959), “Some remarks on changing populations”, *The Kinetics of Cellular Proliferation*, In Grune & Stratton, editor, New York, 382-407.
- [34] Webb G.F., (1985), “Theory of Nonlinear Age-dependent population dynamics”, *Monographs and Textbooks in Pure and Applied mathematics*, Vol 89.